Generalized Boundary Adaptation Rule for minimizing r-th power law distortion in case of high resolution quantization *

Dominique MARTINEZ(†) & **Marc M. VAN HULLE**(‡)

(†) Laboratoire d'Analyse et d'Architecture des Systèmes (LAAS) - CNRS,
7 Av. du Col. Roche, 31077 Toulouse, France

(‡) Laboratorium voor Neuro- en Psychofysiologie,
K.U.Leuven, Campus Gasthuisberg, Herestraat, 3000 Leuven, Belgium

Abstract

A new generalized unsupervised competitive learning rule is introduced for adaptive scalar quantization. The rule, called generalized Boundary Adaptation Rule (BAR_r), minimizes r-th power law distortion D_r in the high resolution case. It is shown by simulations that a fast version of BAR_r outperforms generalized Lloyd I in minimizing D_1 (mean absolute error) and D_2 (mean squared error) distortion with substantially less iterations. In addition, since BAR_r does not require generalized centroid estimation, as in Lloyd I, it is much simpler to implement.

Key words : unsupervised competitive learning, adaptive scalar quantization, high resolution quantization, r-th power law distortion, Lloyd-Max quantizers, generalized Lloyd I, information-theoretic entropy, boundary point estimation.

1 Introduction

A regular N-point scalar quantizer is a function $Q(x)$ which maps a scalar-valued input signal x into one of N quantization levels $y_1, y_2, ..., y_N$. The quantizer is specified by the values of the quantization levels and the N disjoint and exhaustive quantization intervals or partition cells $R_1, R_2, ..., R_N$. Then $Q(x) = y_j$ if $x \in R_j$. A given quantizer is optimal if it minimizes a distortion measure $d(x, Q(x))$, a monotonically increasing, non-negative function of the error magnitude $\mid x - Q(x) \mid$.

Many distortion measures have been proposed in literature (Makhoul *et al.*, 1985) but for regular quantizers, the most commonly used are the Holder norm and its r-th power, the r-th power law distortion:

$$d(x, Q(x)) = D_r \equiv \sum_{i=1}^{N} \int_{x_{i-1}}^{x_i} \mid x - y_i \mid^r p(x) \;\; dx, \tag{1}$$

with $0 < r \leq \infty$ and with $x_0 = -\infty$ and $x_N = \infty$ for a probability density function (*p.d.f.*) $p(x)$ with unbounded support. It is assumed that $p(x)$ has a finite r-th power moment. In quantization, the most commonly used powers are $r = 1$ (mean absolute error) (see *e.g.* Kassam, 1978), $r = 2$ (mean squared error) (see *e.g.* Lloyd, 1957, 1982 and Max, 1960) and $r = \infty$ (mean maximum absolute error).

Lloyd (1957,1982) and Max (1960) independently derived two necessary optimality conditions for $r = 2$ distortion: the centroid and the nearest-neighbor conditions. In case the input distribution is not known, the most widely used design algorithm for scalar quantizers is the (standard) Lloyd I algorithm, and which can be extended for minimizing r-th power law distortion (generalized Lloyd I) (Gersho and Gray, 1991). However, since the input distribution is not known, the computations required to obtain the y_is which satisfy these conditions are iterative and may result in non-optimal quantizers (see *e.g.* Fleisher, 1964). Furthermore, since they are batch algorithms, the design of the quantizer can only begin after the entire

* A version of this paper will appear in the journal Neural Network, 1995.

training set is available. By consequence, these algorithms are not able to accommodate "on-line" changes in the input *p.d.f.*

A number of researchers have developed unsupervised competitive learning algorithms for training Artificial Neural Networks that perform vector and scalar quantization (for references, see Hertz *et al.*, 1991). These algorithms often are neuronal versions of adaptive quantization schemes (Gersho and Gray, 1991) but for which the synaptic weights are updated "on-line." In standard unsupervised competitive learning (standard UCL), the weight updates amount to centroid estimation and nearest-neighbor classification and attempt to minimize $r = 2$ distortion. As pointed out by Grossberg (1976) and Rumelhart and Zipser (1985) among others, one problem with UCL is that some neurons may never win the competition and therefore, never learn (dead units). Although this can be avoided in Kohonen learning by adding a neigborhood to each neuron, due to this neigborhood, D_r with $r = 1/2 + 3/[2(2n+1)^2]$ is minimized instead, with n the number of neighbor neurons on both sides, as shown in (Ritter, 1991) for the scalar case. Therefore, as longs as this neigborhood has not disappeared, D_2 will not be minimized. Dead units also can be avoided if the learning rule attempts to generate neurons that win the competition with equal probability. Several competitive learning rules have been developed around the principle of *equiprobable* quantization (De Sieno, 1988; Van den Bout and Miller, 1989; Ahalt *et al.*, 1990; Van Hulle and Martinez, 1993,1994) but, from the viewpoint of minimizing D_2, equiprobable quantizers are not optimal, or even not near-optimal, as we have proven elsewhere (Van Hulle and Martinez, 1993).

In this article, we introduce a novel unsupervised competitive learning rule, called generalized Boundary Adaptation Rule (BAR_r), for minimizing r-th power law distortion in the case of high resolution quantization. In addition, since we prove that BAR_r converges to a disjunct quantization for which the interval lengths and their probabilities are non-zero for bounded *p.d.f.*s, our learning rule is guaranteed not to produce dead units.

2 High-Resolution Quantization

A regular N point scalar quantizer partitions the real line into N disjoint and exhaustive quantization intervals, *i.e.* the half-open intervals $R_i \equiv [x_{i-1}, x_i)$, $i = 1, ..., N$, with x_{i-1} and x_i the corresponding boundary points, and $x_0 = -\infty < ... < x_{i-1} < x_i < ... < x_N = \infty$. In high-resolution quantizers, the number of quantization intervals N is very large so that the interval lengths are very small and the input *p.d.f.* $p(x)$, which is assumed to be stationary in this article, is roughly constant over the individual intervals. Hence, $p(x) \approx p_i$ in interval R_i, and $p_i = p(R_i)/\delta_i$, with $p(R_i)$ the probability of $x \in R_i$ and with $\delta_i = x_i - x_{i-1}$ the length of R_i. Suppose that $E[x^2] < \infty$ so that with probability nearly one, x takes on values in a finite interval $[a, b)$ hence, since we then can ignore the probability that $x \notin [a, b)$, the r-th power law distortion eq. (1) can be rewritten as:

$$D_r \approx \sum_{i=1}^{N} \frac{p(R_i)}{\delta_i} \int_{x_{i-1}}^{x_i} \mid x - y_i \mid^r \ dx, \tag{2}$$

with $x_0 \equiv a$ and $x_N \equiv b$, and $x_{i-1} < x_i$, $i = 1, ..., N$. The interval $[a, b)$ is called the quantization range. By reformulating the absolute value in eq. (2), one obtains:

$$D_r \approx \sum_{i=1}^{N} \frac{p(R_i)}{\delta_i} [\int_{x_{i-1}}^{y_i} (y_i - x)^r \ dx + \int_{y_i}^{x_i} (x - y_i)^r \ dx]. \tag{3}$$

The high-resolution case enables us to approximate the centroid y_i by the midpoint of the corresponding quantization interval R_i: $y_i \approx \frac{x_{i-1}+x_i}{2}$. Hence, integration of eq. (3) and substitution for the midpoint yields:

$$D_r \approx \frac{1}{2^r(r+1)} \sum_{i=1}^{N} p(R_i) {\delta_i}^r.$$

A necessary condition for minimizing D_r is obtained by the use of the method of Lagrange multipliers, as shown for $r = 2$ in (Panter and Dite, 1951):

$$\delta_j^r \; p(R_j) = \delta_{j+1}^r \; p(R_{j+1}), \quad j = 1, ..., N-1. \tag{4}$$

From eq. (4), it follows that every quantization interval R_i has an identical distortion contribution $D_r(i) = \frac{D_r}{N} \approx \frac{1}{2^r(r+1)} p(R_i) {\delta_i}^r$. This *equidistortion principle* was first observed for the scalar case (Panter and Dite, 1951) and then extended to the vector case (Gersho, 1979). The previous condition can also be interpreted as the solution which maximizes the following information-theoretic entropy measure:

$$I_r = -\sum_{j=1}^{N} \frac{\delta_j^r p(R_j)}{Z} \; log \frac{\delta_j^r p(R_j)}{Z}, \tag{5}$$

with Z the normalization constant:

$$Z = \sum_{j=1}^{N} \delta_j^r p(R_j).$$

As a special case, I_0 corresponds to Shannon entropy and the necessary condition to an equiprobable quantization of $p(x)$. In case $r = \infty$, it corresponds to a uniform quantization. This can be shown as follows.

Proposition 1: The necessary condition eq. (4) for $r = \infty$ corresponds to a uniform quantization of the quantization range $[a, b)$.

Proof:

Consider again eq. (4). Taking the logarithm of both sides yields:

$$r \log \frac{\delta_j}{\delta_{j+1}} = \log \frac{p(R_{j+1})}{p(R_j)},$$

given that $p(R_j) \neq 0$, and $\delta_j^r \neq 0$, $\forall j$. Now since:

$$\lim_{r \to \infty} \log \frac{\delta_j}{\delta_{j+1}} = \lim_{r \to \infty} \frac{1}{r} \log \frac{p(R_{j+1})}{p(R_j)} = 0,$$

we have that $\delta_j = \delta_{j+1}$, $j = 1, ..., N-1$. QED.

Average learning rule

A system of ordinary differential equations (ODEs) which realizes eq. (4) is given by the average learning rule:

$$\frac{dx_j}{dt} = \eta(\delta_{j+1}^r p(R_{j+1}) - \delta_j^r p(R_j)), \quad j = 1, ..., N-1. \tag{6}$$

The equilibrium points of this system are the boundary point vectors $[x_j]$ that satisfy the homogeneous system of equations:

$$0 = \delta_{j+1}^r p(R_{j+1}) - \delta_j^r p(R_j), \quad j = 1, ..., N-1, \tag{7}$$

and thus the necessary condition eq. (4). In Appendix 1, it is proven that the average learning rule converges to a unique boundary point vector $[x_j]$. The assumptions needed to assure a unique convergence are with the input $p.d.f.$ and are slightly different for the $r = 0$ and $r \neq 0$ cases. The lemmas given in Appendix 1 also provide an account on the rule's learning dynamics.

3 Generalized Boundary Adaptation Rule

There are several numerical techniques available to integrate eq. (6) directly, given an initial boundary point vector. However, in case the input $p.d.f.$ is not known *a priori*, the probabilities $p(R_j)$ cannot be determined by solving for the corresponding integrals. Instead, input samples will have to be drawn from a source with $p(x)$ statistics, and boundary points adjusted in response to them. Furthermore, in case the boundary points are updated after the presentation of each input sample, these probabilities will have to be estimated incrementally. The learning rule proposed in this section belongs to this class of parameter estimation techniques.

In order to derive our standard, incremental learning rule, we define $\mathbb{1}_{R_j}(x)$, $j = 1, ..., N$, as the code membership function of interval R_j:

$$\mathbb{1}_{R_j}(x) = \begin{cases} 1 & \text{if } x \in R_j \\ 0 & \text{if } x \notin R_j. \end{cases}$$

The probability $p(R_j)$ of $x \in R_j$ satisfies:

$$p(R_j) = \int_{x_0}^{x_N} \mathbb{1}_{R_j} p(x)\ dx = \int_{x_{j-1}}^{x_j} p(x)\ dx \triangleq E[\mathbb{1}_{R_j}], \tag{8}$$

where $E[.]$ stands for a conditional average over the input $p.d.f.$, given the boundary point vector $[x_j]$.

Assume that for input x, $\mathbb{1}_{R_j} = 1$. We then modify R_j by increasing x_{j-1} and decreasing x_j in proportion to the r-th power of the length of R_j. In its simplest form, the generalized unsupervised learning rule called BAR_r (generalized Boundary Adaptation Rule) reduces to:

$$\Delta x_j = \eta(\delta_{j+1}^r \mathbb{1}_{R_{j+1}} - \delta_j^r \mathbb{1}_{R_j}), \quad j = 1, ..., N-1, \tag{9}$$

with η the learning rate, a positive scalar, and with the $\mathbb{1}_{R_j}$s and δ_js defined with respect to the boundary points at the previous time step. Note that BAR_0 and BAR_1 have previously been introduced in another context (Van Hulle and Martinez, 1993,1994; and Martinez and Van Hulle, 1993, respectively).

Proposition 2: Under the assumption that η is small, BAR_r realizes at convergence the necessary condition eq. (4), minimizes the r-th power law distortion D_r and maximizes I_r.

Proof:

Provided that η is very small, so that the updates Δx_j, $j = 1, ..., N-1$, are much smaller than the interval lengths δ_j, the δ_js will remain positive and correspondingly, the intervals will remain disjunct over time. Furthermore, since the fluctuations around the averages of Δx_j are very small, a small fluctuation expansion will hold so that these averages will approximate the right-hand side of eq. (6). At convergence, we have on average that $E[\Delta x_j] = 0$, $j = 1, ..., N$, and hence:

$$E[\delta_{j+1}^r \mathbb{1}_{R_{j+1}}] = E[\delta_j^r \mathbb{1}_{R_j}], \quad \textit{and}, \tag{10}$$

$$\delta_{j+1}^r E[\mathbb{1}_{R_{j+1}}] = \delta_j^r E[\mathbb{1}_{R_j}], \quad j = 1, ..., N-1 \tag{11}$$

since $E[.]$ is calculated given the x_js at the previous time step. Hence, since $p(R_j) \triangleq E[\mathbb{1}_{R_j}]$, eq. (4) is satisfied and D_r is minimized. By virtue of definition eq. (5), I_r is maximized. QED.

For $r = 0$, we have proven that eq. (9) converges with probability unity to eq. (11) in case the samplings of $p(x)$ are ergodic and stationary (Van Hulle and Martinez, 1994).

A faster rule, called $FBAR_r$, is obtained by updating all boundary points each time an input is presented:

$$\Delta x_j = \eta(\sum_{k=j+1}^{N} \delta_k^r \frac{\mathbb{1}_{R_k}}{N-j} - \sum_{k=1}^{j} \delta_k^r \frac{\mathbb{1}_{R_k}}{j}) \quad j = 1, ..., N-1. \tag{12}$$

The terms in the denominators are needed to assure that $FBAR_r$ satisfies the necessary condition eq. (4) at convergence, as proven in Appendix 2. Since they, in effect, act as asymmetrical connection templates, these terms do not favor a particular shape of input $p.d.f.$ at all. In a previous article (Van Hulle and Martinez, 1993), we have compared the quantization performances of $FBAR_0$ and five popular unsupervised competitive learning rules and the standard Lloyd I algorithm.

Difference between BAR_r and $FBAR_r$

It is customary to view a stochastic process, such as BAR_r or $FBAR_r$, as an average, deterministic trajectory described by the drift vector :

$$DV = [DV_j] \equiv [< \Delta x_j >_\Omega], \tag{13}$$

with stochastic fluctuations around it, described by the diffusion matrix :

$$DM = [DM_{jk}] \equiv < [\Delta x_j][\Delta x_k]' >_{\Xi(t)}, \tag{14}$$

provided that η is very small. The subscript Ω represents an average taken over the stationary $p.d.f.$ $p(x)$ and $\Xi(t)$ an ensemble average. It can be verified that in case $r = 0$, the rate of convergence of DV for a $p.d.f.$ with bounded support and for which the boundary points are initialized outside the $p.d.f.$'s range, is on the order of $\mathcal{O}(\frac{\eta}{N(N-1)})$. As a measure of rate of convergence, we consider that of the most distant boundary point. At convergence, the diagonal components of DM equal $\eta^2 \frac{2}{N}$, in case ergodicity holds.

Under the same conditions as above, the rate of convergence of $FBAR_0$ is now on the order of $\mathcal{O}(\frac{\eta}{N-1})$. At convergence, the diagonal components of DM equal $\eta^2 \frac{1}{(N-j)j}$, so that for N even, $DM_{\frac{N}{2}\frac{N}{2}} = \eta^2 \frac{4}{N^2}$. Hence, for the same N and η, $FBAR_0$ converges N times faster than BAR_0. The stochastic fluctuations around the boundary points of $FBAR_0$ are on the order of N times smaller than those of BAR_0. Furthermore, convergence of $FBAR_0$ and the size of the fluctuations are approximatively independent of the number of quantization intervals provided that η increases in proportion to N, an interesting feature in case of simulations. For $r \neq 0$, these comparisons are more involved since they strongly depend on the initial boundary point vector, the interval lengths, and the $p.d.f.$ at hand. However, also in that case, $FBAR_r$ is expected to be about N times faster than BAR_r, as was confirmed by actual simulations. Hence in case one is interested in determining the equilibrium boundary point vector only, and not the learning dynamics itself, $FBAR_r$ is preferred above BAR_r. For this reason, $FBAR_r$ is used in the next section on simulations.

4 Numerical Comparisons

The most widely used design algorithm for scalar quantizers is the (generalized) Lloyd I algorithm (Lloyd, 1957,1982). This algorithm yields the optimal quantization results in case the input $p.d.f.$s are known *a priori*, and hence it can be used as a reference for comparison. Here, we will compare the performance of $FBAR_r$ with generalized Lloyd I in terms of minimizing D_r for the two most popular cases, $r = 1$ (mean absolute error) and 2 (mean squared error). We will consider both a symmetrical and an asymmetrical input $p.d.f.$ The performance results are shown in Tables 1 and 2 for a Gaussian $p.d.f.$ with mean $\overline{x} = 0$ and standard deviation $\sigma_x = 1$, and Tables 3 and 4 for an exponential $p.d.f.$ with $\overline{x} = 1$. For both tables, the *a priori* and the empirical quantization results are given. The starting configuration is a set of N boundary points drawn randomly from a uniform $p.d.f.$, and labeled in increasing order so that $x_{j-1} < x_j$, $j = 1, ..., N$.

Consider first the *a priori* results: they are obtained using *a priori* knowledge of the Gaussian or exponential $p.d.f.$ The Optimal Quantization results were determined with the generalized Lloyd I algorithm. According to the generalized centroid definition (Gersho and Gray, 1991), the intervals' medians and averages were taken for $r = 1$ and 2, respectively. The optimal High-Resolution Quantization (HRQ) results were determined by solving for the boundary points that both satisfy eq. (4) and minimize eq. (2); the quantization levels correspond to the intervals' midpoints or their centroids (columns "midpoint" and "centroid"). In order to avoid dependency on a particular choice of quantization range, the midpoints for the first and last intervals were taken as follows: $y_1 = x_1 - \frac{x_2 - x_1}{2}$ and $y_N = x_{N-1} + \frac{x_{N-1} - x_{N-2}}{2}$.

For the empirical results, the values given for both generalized Lloyd I and $FBAR_r$ are averages over 20 runs with their standard deviations. In the case of Lloyd I, for each run, the algorithm repeatedly iterated on a

N	*a priori* results			empirical results		
	Optimal Quantiz.	Opt. HRQ		Lloyd I	$FBAR_1$	
		midpoint	centroid		midpoint	centroid
2	0.473	1.710	0.473	0.475 ± 1.9E-3	1.710 ± 6.6E-4	0.474 ± 1.5E-3
4	0.266	0.275	0.271	0.268 ± 2.3E-3	0.305 ± 6.5E-3	0.290 ± 4.1E-3
8	0.143	0.151	0.148	0.146 ± 5.1E-3	0.159 ± 2.2E-3	0.153 ± 1.5E-3
16	0.0745	0.0784	0.0768	0.0821 ± 1.1E-2	0.0789 ± 8.6E-4	0.0773 ± 6.3E-4
32	0.0382	0.0395	0.0390	0.0407 ± 9.9E-3	0.0395 ± 2.8E-4	0.0390 ± 2.0E-4
64	0.0193	0.0197	0.0196	0.0236 ± 2.3E-3	0.0197 ± 9.3E-5	0.0195 ± 6.4E-5
128	0.00972	0.00984	0.00980	0.0136 ± 1.7E-3	0.00983 ± 2.4E-5	0.00980 ± 1.8E-5
256	0.00489	0.00491	0.00490	0.00899 ± 1.2E-3	0.00494 ± 1.3E-5	0.00492 ± 9.0E-6

Table 1: *A priori* and empirical D_1 quantization results as a function of the number of quantization intervals N for a Gaussian *p.d.f.* with $\overline{x} = 0$ and $\sigma_x = 1$. The empirical results are averages over 20 runs with their standard deviations. The results for each run were obtained using the same set of 50,000 samples for both Lloyd I and $FBAR_1$. Lloyd I iterates on a given set of samples until convergence, *i.e.* until the relative variation in D_1 of two subsequent iterations is less than 0.001. $FBAR_1$ iterates only once on that set with $\eta = 10^{-3}N$. For the quantization range $[a, b)$, we took $a = -5$ and $b = 5$.

N	*a priori* results			empirical results		
	Optimal Quantiz.	Opt. HRQ		Lloyd I	$FBAR_2$	
		midpoint	centroid		midpoint	centroid
2	0.363	3.260	0.363	0.365 ± 7.2E-3	3.263 ± 6.7E-3	0.365 ± 2.3E-3
4	0.117	0.180	0.156	0.120 ± 7.8E-3	0.179 ± 6.6E-3	0.156 ± 4.4E-3
8	0.0345	0.0452	0.0411	0.0365 ± 3.9E-3	0.0455 ± 9.5E-4	0.0413 ± 6.1E-4
16	0.00950	0.0109	0.0103	0.0112 ± 1.5E-3	0.0110 ± 1.7E-4	0.0104 ± 1.2E-4
32	0.00250	0.00267	0.00259	0.00383 ± 1.0E-3	0.00268 ± 1.9E-5	0.00260 ± 1.3E-5
64	0.000643	0.000661	0.000652	0.00108 ± 2.7E-4	0.000725 ± 8.0E-6	0.000710 ± 8.0E-6
128	0.000163	0.000165	0.000163	0.000424 ± 1.0E-4	0.000301 ± 5.0E-6	0.000296 ± 5.0E-6
256	0.0000410	0.0000411	0.0000410	0.000289 ± 1.8E-4	0.000163 ± 2.0E-6	0.000161 ± 2.0E-6

Table 2: *A priori* and empirical D_2 quantization results. Same conventions as in Table 1 except that for $FBAR_2$, $\eta = 5\ 10^{-4}N$.

	a priori results			empirical results		
N	Optimal Quantiz.	Opt. HRQ midpoint	Opt. HRQ centroid	Lloyd I	$FBAR_1$ midpoint	$FBAR_1$ centroid
2	0.405	1.205	0.467	0.415 ± 4.1E-2	1.191 ± 4.0E-2	0.471 ± 8.9E-3
4	0.223	0.275	0.244	0.236 ± 5.6E-2	0.273 ± 5.5E-3	0.243 ± 3.1E-3
8	0.118	0.129	0.123	0.120 ± 4.0E-3	0.129 ± 1.5E-3	0.123 ± 1.0E-3
16	0.0606	0.0632	0.0619	0.0622 ± 9.7E-4	0.0635 ± 3.6E-4	0.0621 ± 2.4E-4
32	0.0308	0.0314	0.0311	0.0510 ± 1.2E-2	0.0315 ± 2.0E-4	0.0312 ± 1.6E-4
64	0.0155	0.0156	0.0156	0.0313 ± 4.9E-3	0.0162 ± 1.5E-4	0.0162 ± 1.5E-4
128	0.00778	0.00781	0.00779	0.0195 ± 2.2E-3	0.00936 ± 1.8E-4	0.00937 ± 1.8E-4
256	0.00390	0.00390	0.00390	0.0105 ± 1.0E-3	0.00557 ± 7.0E-6	0.00556 ± 8.0E-6

Table 3: *A priori* and empirical D_1 quantization results as a function of the number of quantization intervals, N, for an exponential $p.d.f.$ with $\overline{x} = 1$. For the empirical results, the values given are averages over 20 runs with their standard deviations. 50,000 samples were used for Lloyd I and for $FBAR_1$ with $\eta = 10^{-3}N$. For the quantization range $[a, b)$, we took $a = 0$ and $b = 15$.

	a priori results			empirical results		
N	Optimal Quantiz.	Opt. HRQ midpoint	Opt. HRQ centroid	Lloyd I	$FBAR_2$ midpoint	$FBAR_2$ centroid
2	0.352	2.268	0.512	0.353 ± 1.6E-3	2.254 ± 1.9E-1	0.525 ± 3.7E-2
4	0.109	0.182	0.134	0.113 ± 6.2E-3	0.186 ± 9.5E-3	0.137 ± 5.6E-3
8	0.0307	0.0376	0.0338	0.0346 ± 6.8E-3	0.0379 ± 1.1E-3	0.0338 ± 7.3E-4
16	0.00819	0.00889	0.00849	0.0137 ± 4.7E-3	0.00898 ± 2.1E-4	0.00854 ± 1.8E-4
32	0.00211	0.00218	0.00213	0.00566 ± 2.0E-3	0.00259 ± 1.6E-4	0.00256 ± 1.7E-4
64	0.000535	0.000541	0.000536	0.00198 ± 6.1E-4	0.00104 ± 7.6E-5	0.00103 ± 7.8E-5
128	0.000134	0.000135	0.000134	0.000694 ± 3.8E-4	0.000417 ± 3.1E-5	0.000421 ± 2.8E-5
256	0.0000336	0.0000337	0.0000336	0.000222 ± 4.2E-5	0.000169 ± 1.1E-5	0.000169 ± 1.1E-5

Table 4: *A priori* and empirical D_2 quantization results. Same conventions as in Table 3 except that for $FBAR_2$, $\eta = 5\ 10^{-4}N$.

given set of 50,000 samples until convergence: this way, for $N = 32$ about 50 iterations, and for $N = 256$ about 140 iterations on average were needed for each run in case of a Gaussian. For $FBAR_r$, in each run, the same set of samples was used but only once for $r = 1$ and twice for $r = 2$. For both generalized Lloyd I and $FBAR_r$, the D_r values were estimated after convergence of the respective algorithms using 100,000 samples. For $FBAR_r$, D_r was calculated using two different quantization levels: midpoints or centroids (columns "midpoint" and "centroid"); for generalized Lloyd I only centroids were used, since these values are directly produced by the algorithm.

We see that, as N increases, the empirical midpoint- and centroid results for $FBAR_r$ approximate well their respective optimal HRQ results, for both $r = 1$ and 2, and that there is no difference in performance for a symmetrical or an asymmetrical *p.d.f.* Furthermore, we observe that $FBAR_r$ ("centroid" column) outperforms empirical Lloyd I for the Gaussian and the exponential *p.d.f* when $N \geq 16$ for $r = 1$ or $r = 2$. In addition, there is considerable variation in the empirical Lloyd I results, compared to $FBAR_r$. Note that for small N, a better $FBAR_r$ result is obtained by taking the intervals' centroids as their quantization levels in the estimation of D_r instead of the midpoints. However for larger Ns, the difference becomes negligible, as expected (*cf.* high resolution assumption).

5 Conclusion

In this article, an unsupervised competitive learning rule, called the generalized boundary adaptation rule, BAR_r, has been introduced for minimizing r-th power law distortion D_r of univariate *p.d.f.*s in the high resolution case. It has been shown that for large N, a fast version of this rule, called $FBAR_r$, is superior to the generalized Lloyd I algorithm if $r = 1$ and 2, for a Gaussian and an exponential *p.d.f.* In addition, unlike Lloyd I, the generalized boundary adaptation rule adapts the quantization intervals "on the fly", *i.e.* after the presentation of each input sample, and in this way, on-line changes in the input *p.d.f.* can be accommodated. Finally, for $r \neq 1$ or 2, the calculation of the generalized centroids in Lloyd I is cumbersome since it requires the use of convex programming techniques (Linde *et al.*, 1980). Instead, BAR_r and $FBAR_r$ do not require centroid estimation and hence, are much simpler to implement.

References

Ahalt, S.C., Krishnamurthy, A.K., Chen, P., & Melton, D.E. (1990). Competitive learning algorithms for vector quantization. *Neural Networks*, vol. 3, 277-290.

DeSieno, D. (1988). Adding a conscience to competitive learning. *Proc. 1988 Int. Conf. Neural Networks (ICNN-88)*, San Diego, vol. 1, 117-124.

Fleisher, P.E. (1964). Sufficient Conditions for Achieving Minimum Distortion in a Quantizer. *IEEE Int. Conv. Rec.*, Part I, 104-111.

Gersho, A. (1979). Asymptotically optimal block quantization. *IEEE Trans. Inform. Theory*, **IT-25**, 373-380.

Gersho, A., & Gray, R.M. (1991). *Vector quantization and signal compression.* Boston, Dordrecht, London: Kluwer.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding. *Biol. Cybern.*, vol 23, 121-134.

Hertz, J., Krogh, A., & Palmer, R.G. (1991). *Introduction to the theory of neural computation.* Reading MA: Addison-Wesley.

Kassam, S.A. (1978). Quantization based on the mean-absolute-error criterion. *IEEE Trans. on Communications*, **COM-26**, 267-270.

Linde, Y., Buzo, A., & Gray, R.M. (1980). An Algorithm for Vector Quantizer Design. *IEEE Trans. on Communications*, **COM-28**, 84-95.

Lloyd, S.P. (1957). *Least squares quantization in PCM's.* Bell Telephone Laboratories Paper, Murray Hill, NJ.

Lloyd, S.P. (1982). Least squares quantization in PCM. *IEEE Trans. Inform. Theory*, **IT-28**, 127-135.

Makhoul, J., Roucos, S., & Gish, H. (1985). Vector quantization in speech coding. *Proceedings of the IEEE*, vol. 73, **11**, 1551-1588.

Martinez, D., & Van Hulle, M.M. (1993). Recurrent Neural Network for Adaptive Non-uniform A/D Conversion. *Proc. World Congress on Neural Networks (WCNN'93)*, Portland, Oregon, Vol. 4, 576-579.

Max, J. (1960). Quantizing for minimum distortion. *IRE Trans. on Inform. Theory*, **IT-6**, 7-12.

Panter, P.F. & Dite, W. (1951). Quantization distortion in Pulse-Count Modulation with nonuniform spacing of levels. *Proc. IRE*, **39**, 44-48.

Ritter, H. (1991). Asymptotic level density for a class of vector quantization processes. *IEEE Transactions on Neural Networks*, vol. 2, no.1, 173-175.

Rumelhart, D.E., & Zipser, D. (1985). Feature discovery by competitive learning. *Cognitive Sci.*, vol. 9, 75-112.

Van den Bout, D. E., & Miller III, T.K. (1989). TInMANN: The integer Markovian artificial neural network. *Proc. Int. Joint Conf. Neural Networks*, Englewood Cliffs, NJ:Erlbaum, II205-II211.

Van Hulle, M.M., & Martinez, D. (1993). On an unsupervised learning rule for scalar quantization following the maximum entropy principle. *Neural Computation*, **5**, 939-953.

Van Hulle, M.M., & Martinez, D. (1994). On a novel unsupervised competitive learning algorithm for scalar quantization. *IEEE Transactions on Neural Networks*, **5**, 498-501.

Appendix 1

The average learning rule eq. (6) converges to a unique boundary point vector $[x_i]$ which satisfies the necessary condition eq. (4). We will show this in two steps. First, we will derive the conditions for which the system of equations eq. (7) has only one non-trivial solution. Then, we will show that the average learning rule has a Liapunov function. Hence, the average learning rule converges to this unique boundary point vector.

Since eq. (7) is a homogeneous system of linear equations with less equations than unknowns, *i.e.* the variables $\delta_k^r p(R_k)$, $k = 1, \ldots N$, we have the trivial solution and a set of non-trivial solutions. We now show that the trivial solution cannot occur. In the following, we will assume that the input *p.d.f.*s are bounded ($p(x) \neq \infty$) so that the corresponding interval lengths are not infinitesimally small. The *support* of the input *p.d.f.*s may be bounded or unbounded (x takes on values in a finite or infinite interval, respectively). We further assume that the boundary points are arranged in increasing order, $x_0 \leq x_j < x_N$, $\forall j$.

Lemma 1: The homogeneous system of equations eq. (7) has only non-trivial solutions for bounded *p.d.f.*s.

Proof:

The trivial solution, $\delta_k^r p(R_k) = 0$, $\forall k$, does not occur since at convergence we have that: a) the quantization range, $b - a$, equals $\sum_{j=1}^N \delta_j$, a non-zero quantity, hence $\sum_{j=1}^N \delta_j^r$, is a non-zero quantity, and b) $\sum_{j=1}^N p(R_j) = 1$ with $p(R_j)$ non-negative, $\forall j$. Furthermore since $p(R_j) = \int_{R_j} p(x)\, dx$, activations are coupled to interval lengths, we conclude that the trivial solution is not feasible. QED.

Lemma 2: The homogeneous system eq. (7) with $r \neq 0$ has, in the case of a bounded *p.d.f.* $p(x)$ with bounded support, only one boundary point vector as non-trivial solution.

Proof by indirect demonstration:

We assume that there are two solutions $Q_1 = \delta_k^r p(R_k)$ and $Q_2 = \delta_k^r p(R_k)$, $k = 1, \ldots N$, with different values for the boundary points. We know from lemma 1 that the trivial solution is not feasible.

Assume that $Q_1 < Q_2$ and that outside interval $[x_0, x_N)$, $x_0 < x_N$, the *p.d.f.* $p(x) = 0$. Assume further that we dispose of the boundary points that realize $Q_1 = \delta_k^r p(R_k)$, $k = 1, \ldots N$. We will now try to find those that realize Q_2. Consider the first quantization interval $R_1 = [x_0, x_1)$ and $Q_1 = \delta_1^r p(R_1)$. In order for the right hand side of the previous equality to become equal to Q_2, x_1 will have to increase (x_0 is fixed), since δ_1^r is a strictly monotonically increasing function and $p(R_1)$ a monotonically increasing function of the interval length and thus of x_1. However, when x_1 increases, the length of R_2 will decrease, and hence x_2 will have to increase in order to achieve $Q_2 = \delta_2^r p(R_2)$ for the same reason as above, and so on until we have reached the last interval R_N. Now since x_N is fixed, the length of R_N is decreased, and $\delta_N^r p(R_N)$ has become smaller than Q_1, we will never be able to attain Q_2.

We can repeat the same reasoning in the case of $Q_1 > Q_2$ to arrive at the same conclusion.

As a consequence, the foregoing is feasible only when $Q_1 = Q_2$, and since $\delta_k^r p(R_k)$, $k = 1, ...N$ is a strictly monotonically increasing function of the interval length, we have that only a single boundary point vector can be the solution of eq. (7). QED.

Strictly speaking, the previous lemma does not hold for a *p.d.f.* with unbounded support. However, we can approximate the bounded support case by choosing the range $[x_0, x_N)$ in such a way that the probability that x belongs to it is arbitrarily close to unity.

In case $r = 0$, the necessary condition eq. (4) implies that $p(R_1) = p(R_2) = ... = p(R_N) = \frac{1}{N}$. When the *p.d.f.* comprises a set of unconnected modes or clusters, separated by domains where $p(x) = 0$, then the possibility exists that there will be an infinite number of valid boundary point vector solutions. However, under a fairly conservative condition, we can assure that there will again be a unique boundary point vector solution. In contrast with the $r \neq 0$ case, the *p.d.f.* may be one with unbounded support.

Lemma 2bis: Let $\Xi = \{x_1, x_2, ..., x_{N-1}\}$ be a boundary point vector that is a solution of the system of equations eq. (7) with $r = 0$. In case for all boundary points $x_j \in \Xi$ holds that $p(x_j) \neq 0$, or when $p(x_j) = 0$ the left- and righthand derivatives in x_j are non-zero, with $p(.)$ a bounded *p.d.f.* with bounded or unbounded support, then Ξ will be the unique solution.

Proof:

The proof is analogous to that of lemma 2 since $p(R_j)$, $j = 1, ..., N-1$ is a strictly monotonously increasing function of the interval length around the equilibrium point $\Xi = \{x_1, x_2, ..., x_{N-1}\}$. QED.

Note that notwithstanding the boundary points may converge to domains where $p(x) = 0$, the necessary condition eq. (4) will always be satisfied. Strictly speaking, such a boundary point vector represents a valid solution, but it is not the only one. This is because eq. (4) does not contain any dependency on the input x or *e.g.* the centroids of the quantization intervals. An additional criterion would, in that case, be required to arrive at a unique boundary point vector solution.

In lemmas 2 and 2bis, the conditions were given for which the average learning rule eq. (6) possesses a unique equilibrium point. Let Q be the value adopted by the equalities in eq. (4). We introduce the following definitions.

<u>Definitions:</u>

Define C_j^2, with $C_j = \delta_j^r p(R_j) - Q$, as a measure for how far interval R_j is away from its equilibrium solution Q. Define for the entire quantizer the set $C = \{C_1, ..., C_N\}$; R_j is called a "maximal" interval when C_j^2 is a maximum of the set of squared elements of C. Define $RB = \{R_j, R_{j+1}, ..., R_{j+(m-1)}\}$ as a block of m subsequent maximal intervals, $m \geq 1$, for which the corresponding $C_j, C_{j+1}, ..., C_{j+(m-1)}$ have the same sign (Fig. 1A). In case $m = 1$, the only interval is termed an "isolated" maximal interval.

We will now show that a Liapunov function $V(C) = \max_{C_j \in C}\{C_j^2\}$ exists which states that the magnitude of the largest squared elements of C decrease over time. Before we can show this, we still need to prove some additional lemmas: the possibility exists that $\frac{dV}{dt} = 0$ occurs without having that $V = 0$. There are two reasons for this:

1. Assume that we have multiple maximal intervals and that three or more of them form a block RB. Without loss of generality we assume that RB comprises three maximal intervals: R_{j-1}, R_j and R_{j+1} (*e.g.* RB_2 in Fig. 1A). Now since $C_{j-1} = C_j = C_{j+1}$, we have that $\frac{dx_{j-1}}{dt} = 0$ and $\frac{dx_j}{dt} = 0$. Hence, C_j does not change so that $\frac{dV}{dt} = 0$.
2. Only for $r = 0$. Assume that R_i is an isolated maximal interval: C_i is a maximal or minimal term in C

and terms C_{i-1} and C_{i+1} are not. In that case, it could happen that both $p(x_{i-1}) = 0$ and $p(x_i) = 0$ so that $\frac{dp(R_i)}{dt} = 0$, and thus that $\frac{dV}{dt} = 0$. The same reasoning holds for the first and last intervals of a block of maximal intervals.

We will now show that $\frac{dV}{dt} = 0$ can occur only during maximally a finite amount of time. After that, V will again be strictly monotonically decreasing, $\frac{dV}{dt} < 0$. We will proceed in a number of steps (see also Fig. 1A,B).

Lemma 3: Let RB be a block of $m > 1$ subsequent maximal intervals for which $\frac{dV}{dt} = 0$. The input *p.d.f.* is as defined in lemmas 2 and 2bis. In case $r \neq 0$, we have for the first interval R_j of RB $(j > 1)$ that $\frac{dC_j^2}{dt} < 0$ and for the last interval R_k of RB $(k = j + (m-1) < N)$ that $\frac{dC_k^2}{dt} < 0$. In case $r = 0$, we have for these intervals j, k that $\frac{dC_{j,k}^2}{dt} \leq 0$ with $\frac{dC_{j,k}^2}{dt} = 0$ only during maximally a finite amount of time.

Proof:

Case $r \neq 0$

Consider the first interval $R_j \in RB$ $(j > 1)$. We have that:

$$\frac{dC_j^2}{dt} = 2[\delta_j^r p(R_j) - Q][\frac{\partial}{\partial x_j}(\delta_j^r p(R_j))\frac{dx_j}{dt} + \frac{\partial}{\partial x_{j-1}}(\delta_j^r p(R_j))\frac{dx_{j-1}}{dt}], \tag{15}$$

with $\frac{\partial}{\partial x_j}(\delta_j^r p(R_j)) > 0$ and $\frac{\partial}{\partial x_{j-1}}(\delta_j^r p(R_j)) < 0$, after some algebraic manipulations. Assume that $\delta_j^r p(R_j) - Q > 0$, which in fact implies that $\delta_j^r p(R_j)$ is the largest term in the ODE of x_{j-1}. Hence, we have that $\frac{dx_{j-1}}{dt} > 0$. We know that $\frac{dx_j}{dt} = 0$. From the previous inequalities we conclude that $\frac{dC_j^2}{dt} < 0$. A similar reasoning for $\delta_j^r p(R_j) - Q < 0$ leads to the same conclusion.

In the same way, we obtain for the last interval $R_{j+(m-1)}$ $(j + (m-1) < N)$ of RB that $\frac{dC_{j+(m-1)}^2}{dt} < 0$.

Case $r = 0$

Consider the first interval $R_j \in RB$. Since R_{j-1} is not a maximal interval, $\frac{dx_{j-1}}{dt} \neq 0$. Note that $\frac{dC_j^2}{dt} = 0$ occurs only when $p(x_{j-1}) = 0$. Since R_j is a maximal interval, we can consider two cases. Firstly, assume that $p(R_j) > Q$: in that case $\frac{dx_{j-1}}{dt} > 0$. However since $p(R_j) \neq 0$ and the input *p.d.f.* is bounded, we can have only a finite distance over which x_{j-1} can increase until $p(x_{j-1}) \neq 0$ and $\frac{dp(R_j)}{dt} \neq 0$, and thus $\frac{dC_j^2}{dt} < 0$. Since over this distance $\frac{dx_{j-1}}{dt}$ remains constant, it will take maximally a finite amount of time until we have that $\frac{dC_j^2}{dt} < 0$.

Secondly, in case $p(R_j) < Q$, we have that $\frac{dx_{j-1}}{dt} < 0$. Since $p(R_{j-1}) \neq 0$ (R_{j-1} is not a maximal interval), a similar reasoning leads to the same conclusion that it will take maximally a finite amount of time until we have that $\frac{dC_j^2}{dt} < 0$ again.

A similar reasoning holds for the last interval $R_{j+(m-1)}$. QED.

Lemma 3bis: Let $\{RB_1, RB_2, ...\}$ be the set of isolated blocks of maximal intervals with equal C terms and with $\frac{dV}{dt} = 0$. In that case, each of these blocks will disappear after maximally a finite amount of time, after which we will have again that $\frac{dV}{dt} < 0$.

Proof:

Without loss of generality, we assume that we have only 1 isolated block. Then, by repeatedly applying lemma 3, we conclude that after maximally a finite amount of time $\frac{dV}{dt} < 0$ again. QED.

Lemma 4: $V(C) = \max_{C_j \in C}\{C_j^2\}$ is a Liapunov function with $\frac{dV}{dt} \leq 0$ and for which $\frac{dV}{dt} = 0$ can occur only during maximally a finite amount of time, unless $\frac{dC_j^2}{dt} = 0, \forall j$ and the equilibrium point is reached.

Proof:

¿From the definition of the terms C_j and the input $p.d.f.$ follows that $V(C)$ and the partial derivatives of the latter are continuous. In addition, $V(C)$ is positive definite, *i.e.* $V(C=\{0,...,0\})=0$ and $V(C)>0$ when $C \neq \{0,...,0\}$. We will now show that the derivative $\frac{dV}{dt}$ is negative semidefinite, namely that $\frac{dV(C \neq \{0,...,0\})}{dt} \leq 0$ and $\frac{dV(C=\{0,...,0\})}{dt}=0$.

We assume first that there is only 1 maximal interval, R_j. If we calculate the derivative of C_j^2, then we obtain:

$$\frac{dC_j^2}{dt} = 2[\delta_j^r p(R_j) - Q][\frac{\partial}{\partial x_j}(\delta_j^r p(R_j))\frac{dx_j}{dt} + \frac{\partial}{\partial x_{j-1}}(\delta_j^r p(R_j))\frac{dx_{j-1}}{dt}]. \tag{16}$$

In case $r \neq 0$, we have that $\frac{\partial}{\partial x_j}(\delta_j^r p(R_j)) > 0$ and $\frac{\partial}{\partial x_{j-1}}(\delta_j^r p(R_j)) < 0$, after some algebraic manipulations; in case $r=0$, we have that $\frac{\partial}{\partial x_j}p(R_j) \geq 0$ and $\frac{\partial}{\partial x_{j-1}}p(R_j) \leq 0$. We now can distinguish two cases. First, $\delta_j^r p(R_j) - Q > 0$, which signifies that $\delta_j^r p(R_j)$ is the largest term in the ODE eq. (6). Hence, we have that $\frac{dx_j}{dt} < 0$ and $\frac{dx_{j-1}}{dt} > 0$. Based on the previous inequalities, we conclude that $\frac{dC_j^2}{dt} < 0$ for $r \neq 0$ and $\frac{dC_j^2}{dt} \leq 0$ for $r=0$. For the latter one, we know from lemma 3 that $\frac{dC_j^2}{dt}=0$ occurs during maximally a finite amount of time only.

Second, we can have that $\delta_j^r p(R_j) - Q < 0$. Then, again based on the corresponding inequalities, we conclude that $\frac{dC_j^2}{dt} < 0$ for $r \neq 0$ and $\frac{dC_j^2}{dt} \leq 0$ for $r=0$. For the latter one, in case $\frac{dC_j^2}{dt}=0$, we re-apply lemma 3.

Assume now that there are several blocks RB of maximal intervals with equal C terms and that $\frac{dV}{dt}=0$. We know from lemma 3bis that none of these blocks is a stable configuration and that after maximally a finite amount of time, we will have again that $\frac{dV}{dt} < 0$.

It is possible that after some time, several blocks RB emerge for which $\frac{dV}{dt}=0$. In that case, we again apply the previous reasoning. The emergence of several blocks RB with equal C terms is not a stable configuration, except when the first term between square brackets in eq. (16) equals zero for all intervals. In that case, $\frac{dC_j^2}{dt}=0$, $j=1,...,N$, and thus $\frac{dV}{dt}=0$ and $\frac{dx_j}{dt}=0$, $j=1,...,N-1$. However, we then also have that $V=0$, so that we have reached the equilibrium point. QED.

We can also show that for $r \neq 0$, V is radially unbounded so that convergence will hold "in the large": when $x_j \to \infty \Rightarrow V \to \infty$, $j=1,...,N-1$. For $r=0$ holds that V is radially bounded: $x_j \to \infty \Rightarrow V \to V_{max} \leq (1-\frac{1}{N})^2$.

Theorem 1: The N-point scalar quantizer with average learning rule eq. (6) converges to a unique boundary point vector under the conditions mentioned in lemmas 2 and 2bis.

Proof:

Under the conditions mentioned in lemmas 2 and 2bis, we have only a single Q, the value adopted by the equalities in eq. (4), and only a single corresponding boundary point vector solution. In lemma 4, we have shown that a Liapunov function exists which converges towards Q. Hence, we have that the average learning rule eq. (6) converges to a unique boundary point vector which satisfies the equalities in eq. (4). QED.

It can further be shown that convergence also holds when we start from a non-ordered set of boundary points provided that $\mathbb{1}_{R_j}$ equals unity when the input x lies somewhere between x_j and x_{j-1}, $\forall j$.

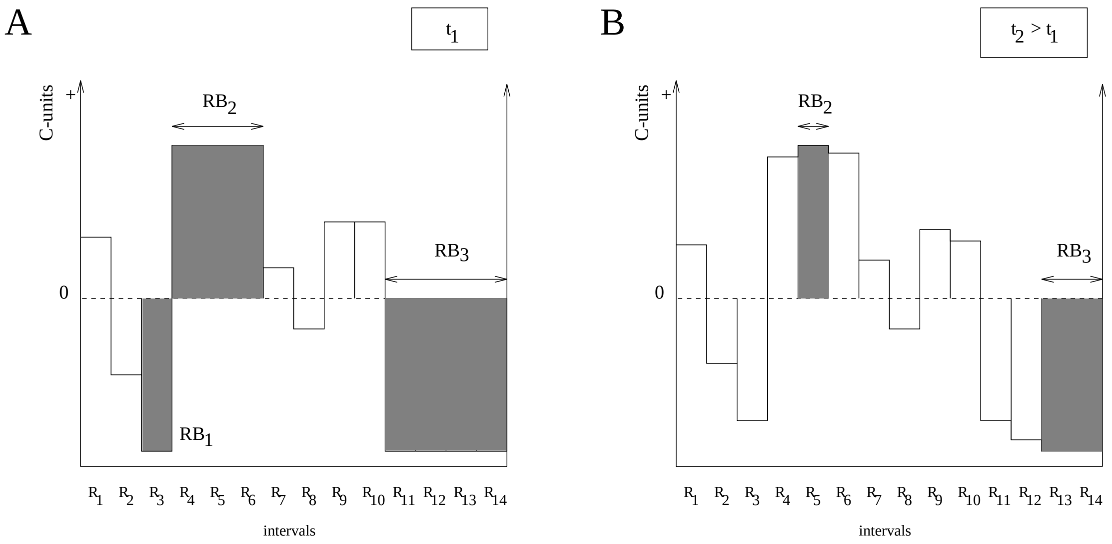

Figure 1: Hypothetical case of a 14-point scalar quantizer where $\frac{dV}{dt} = 0$. To every interval corresponds a positive or negative C-value. The shaded areas denote "maximal" intervals, *i.e.* intervals whose squared C-values are maximal. A series of neighboring maximal intervals for which the corresponding C-values have the same sign, is called a block (RB). (A) In this quantizer we have at time t_1, 3 blocks of maximal intervals: RB_1, RB_2 and RB_3. Note that since RB_1 contains only a single interval, it is termed an "isolated" maximal interval. (B) At time $t_2 > t_1$ convergence has progressed: RB_1 has disappeared; RB_2 became a isolated maximal interval; and RB_3 now comprises only 2 intervals instead of 4.

Appendix 2

Theorem 2: $FBAR_r$ eq. (12) satisfies the necessary condition eq. (4) at convergence.

Proof:

At convergence, we have that $E[\Delta x_j] = 0, j = 1, ..., N-1$, hence eq. (12) becomes:

$$\sum_{k=j+1}^{N} \delta_k^r \frac{E[\mathbb{1}_{R_k}^t]}{N-j} - \sum_{k=1}^{j} \delta_k^r \frac{E[\mathbb{1}_{R_k}^t]}{j} = 0 \quad j = 1, ..., N-1. \tag{17}$$

Since eq. (17) is a homogeneous system of linear equations with fewer equations than unknowns, *i.e.* the $\delta_k^r E[\mathbb{1}_{R_k}^t]$, $k = 1, ...N$, we have the trivial solution and a set of non-trivial solutions. We now show that the trivial solution cannot occur and the non-trivial solutions satisfy the necessary condition eq. (4).

1) The trivial solution, $\delta_k^r E[\mathbb{1}_{R_k}^t] = 0$, $\forall k$, does not occur since at convergence we have that: a) the quantization range, $b - a$, equals $\sum_{j=1}^{N} \delta_j$, a non-zero quantity, hence $\sum_{j=1}^{N} \delta_j^r$, is a non-zero quantity, and b) $\sum_{j=1}^{N} E[\mathbb{1}_{R_j}^t] = 1$ with $E[\mathbb{1}_{R_j}^t]$ non-negative, $\forall j$. Furthermore since $p(R_j) \triangleq E[\mathbb{1}_{R_j}]$, $\forall j$, and since $p(R_j) = \int_{R_j} p(x)\, dx$, activations are coupled to interval lenghts, we conclude that the trivial solution is not feasible.

2) Consider the necessary condition eq. (4). Assume that we multiply the j-th equation of eq. (17) with $N - j$, $\forall j$, so that we obtain a new system of equations. We now can find eq. (4) recursively, one equality at a time. We start with the first two equations of the new system and subtract the second from the first. We then obtain that $\delta_1^r E[\mathbb{1}_{R_1}^t] = \delta_2^r E[\mathbb{1}_{R_2}^t]$. By subtracting the third from the second and substituting the equality just found, we obtain that $\delta_2^r E[\mathbb{1}_{R_2}^t] = \delta_3^r E[\mathbb{1}_{R_3}^t]$, and so on until we have reached the last two equations. Now since $p(R_j) \triangleq E[\mathbb{1}_{R_j}]$, $\forall j$, we then have that eq. (4) is satisfied. QED.

Nomenclature

$\mathbb{1}_{R_j}$: code membership function of interval R_j

BAR_r : generalized Boundary Adaptation Rule, eq. (9)

D_r : r-th power law distortion, eq. (2)

$d(x, Q(x))$: distortion metric defined between input x and quantized output $Q(x)$, eq. (1)

δ_i : length of interval R_i

η : learning rate

$FBAR_r$: generalized Fast Boundary Adaptation Rule, eq. (12)

HRQ : High-Resolution Quantization

I_r : (generalized) information-theoretic entropy measure, eq. (5)

N : number of quantization levels

$Q(x)$: function which maps input x onto a discrete number of quantization levels (formalizes scalar quantizer)

$p(x)$ probability density function of x

R_i : i-th quantization interval or partition cell

x_i : i-th boundary point (x_{i-1} and x_i define R_i)

y_i : i-th quantization level

Z : normalization constant (*cf.* I_r)