AD-A120 260 TEXAS UNIV AT AUSTIN DEPT OF ELECTRICAL ENGINEERING F/G 12/1
SOME REMARKS ON OPTIMAL QUANTIZATION.(U)
JUL 82 E F ABAYA, G L WISE AFOSR-81-0047
UNCLASSIFIED AFOSR-TR-82-0820 NL

1 OF 1
AD A
120260

END
DATE
FILMED
11-82
DTIC

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(2)

AD A120260

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR- 82-0820
2. GOVT ACCESSION NO.: AD-A120 260
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): SOME REMARKS ON OPTIMAL QUANTIZATION
5. TYPE OF REPORT & PERIOD COVERED: TECHNICAL
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Efren F. Abaya and Gary L. Wise
8. CONTRACT OR GRANT NUMBER(s): AFOSR-81-0047
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of Electrical Engineering, University of Texas, Austin TX 78712
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: PE61102F; 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS: Mathematical & Information Sciences Directorate, Air Force Office of Scientific Research, Bolling AFB DC 20332
12. REPORT DATE: 8 JUL 82
13. NUMBER OF PAGES: 6
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:
16. DISTRIBUTION STATEMENT (of this Report): Approved for public release; distribution unlimited.

DTIC ELECTE OCT 14 1982 E

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):
18. SUPPLEMENTARY NOTES: Presented at the 1982 Conference on Information Sciences and Systems, March 17-19, 1982, Princeton, New Jersey. To be published in the Proceedings of the Conference.
19. KEY WORDS (Continue on reverse side if necessary and identify by block number): Existence of optimal quantizers, uniqueness of quantizers, symmetry of quantizers, convergence properties of quantizers.
20. ABSTRACT (Continue on reverse side if necessary and identify by block number):

Current numerical algorithms for constructing optimal quantizers encounter difficulties if either an optimal solution does not exist, or if local minima are present. In this paper, the authors show that an optimal quantizer may not always exist, and present a condition on the distortion measure, and another condition on the probability distribution, either of which guarantees existence of an optimal solution. Then we comment on the role of symmetry in optimal quantization and discuss the validity of one method that has been used to prove uniqueness of the optimal solution.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DTIC FILE COPY

82 10 12 201

AFOSR-TR- 82-0820

SOME REMARKS ON OPTIMAL QUANTIZATION

Efren F. Abaya
Department of Electrical Engineering
University of Texas at Austin
Austin, Texas 78712

and

Gary L. Wise
Departments of Electrical Engineering and Mathematics
University of Texas at Austin
Austin, Texas 78712

AFOSR-81-0047

ABSTRACT

Current numerical algorithms for constructing optimal quantizers encounter difficulties if either an optimal solution does not exist, or if local minima are present. In this paper we show that an optimal quantizer may not always exist, and present a condition on the distortion measure, and another condition on the probability distribution, either of which guarantees existence of an optimal solution. Then we comment on the role of symmetry in optimal quantization and discuss the validity of one method that has been used to prove uniqueness of the optimal solution.

I. Introduction and Preliminaries

In recent years much work has been done to investigate design algorithms for optimal quantization [1-5]. A popular approach is to use numerical techniques to minimize a mean distortion function. With this philosophy, one way to generate an optimal design is to successively improve on a sequence of sub-optimal designs using conditions which are necessary (but not sufficient) for a stationary point. Max [1] in 1960 developed a set of necessary conditions applicable to differentiable cost functions, and suggested one algorithm for the mean square distortion measure. Variants of this algorithm have been successfully applied to scalar quantization problems [1,4]. Linde, et al. [2], generalizing on the work of Lloyd [10], presented a set of necessary conditions similar to those of Max, but having much wider applicability. Their algorithm handles arbitrary multi-dimensional distributions (i.e., vector quantizers) and very general distortion measures. Two considerations arise in the use of either of the above algorithms. The first question is the existence of an optimum quantizer (i.e., a solution to the optimization problem). This is directly related to the convergence of the two algorithms, because if an optimum does not exist, then the algorithms either do not converge, or may converge to a non-minimum solution. We will present an example of the latter possibility. It is of interest, therefore, to set down some conditions under which a global minimum may exist. For twice-differentiable probability density functions an easily checked convexity test is available [6]. Essentially this determines when the optimization problem has a unique stationary point. However, there are many probability densities of practical interest for which the criterion does not apply. A different viewpoint was taken in [7,8] which presented sufficient conditions on the form of the distortion function. Both of these results are summarized in the development.

The second question relates to the nature of the solutions obtained by the above algorithms. It has been demonstrated that these algorithms can sometimes converge to a quantizer which is locally optimum, but does not have the globally minimum distortion [2]. This can obviously present some difficulties in problems with several minima. Some modifications in the optimization algorithms to overcome this problem have been suggested [2,9], but this greatly increases the computation time. A criterion to determine the uniqueness of stationary points would be useful. For twice-differentiable density functions the existence criterion in [6] is applicable. In general, however, determining uniqueness seems to be a more difficult problem than the first. The authors are not aware of any other results which address the problem of uniqueness.

Recent algorithms already are capable of handling general distributions and complex cost functions [2] and it will be correspondingly more difficult tc answer the two questions presented above. Many of the previously mentioned results may not be applicable anymore. Therefore these questions need to be examined further so that existing optimization procedures may be used with confidence that the resulting solution is indeed optimal.

In this paper, we first develop necessary and sufficient conditions for the existence of multidimensional quantizers. Then we turn to the question of uniqueness and, through several examples, point out some interesting results.

First we give a few definitions and establish some notation. An N-level k-dimensional vector quantizer is a mapping Q: $\mathbb{R}^k \to \mathbb{R}^k$ which assigns to the input vector x an output vector Q(x) chosen from a finite set of N distinct vectors $\{y_i: y_i \in \mathbb{R}^k, i=1,2,\ldots,N\}$. When optimal quantizers are being considered with an error based on the Euclidean norm, there is no loss of generality in assuming the "nearest neighbor" assignment rule: Q(x) is that member of the set which is nearest to x in Euclidean norm, with ties being broken in some pre-assigned manner. This rule will be adopted throughout the rest of this paper. Scalar quantization (k=1) is considered to be a special case of vector quantization. Let X be a random vector taking values in $\mathbb{R}^k$ and having a cumulative distribution function F. A measure of the performance of a quantizer Q applied to the random vector X is the mean distortion function

$$D = D(Q,F) = \int C_o(\|x - Q(x)\|)dF(x). \quad (1)$$

Here $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^k$ and $C_o(t)$ is an appropriate cost function. For example, the

Presented at the 1982 Conference on Information Sciences and Systems, March 17-19, 1982; to be published in the Proceedings of the conference.

82 12 201

r^{th}-power cost function $C_o(t) = t^r$ is a popular choice. For simplicity we write $C_o(\|x\|) = C(x)$.

Because of nearest neighbor assignment, a quantizer is fully specified by its set of output levels $\{y_i\}$. Thus when F is fixed we may write $D(Q,F) = D(y_1,y_2,\ldots,y_N)$. The quantizer optimization problem is to find the global minimum point of $D(y_1,y_2,\ldots,y_N)$, if any. A quantizer is said to be locally optimum if its output levels form a locally optimum point of $D(y_1,y_2,\ldots,y_N)$.

Let Q_n and Q be vector quantizers. We say that the sequence $\{Q_n\}$ converges to Q if $Q_n(x) \to Q(x)$ at all continuity points of Q. Equivalently, in view of nearest neighbor assignment, $\{Q_n\}$ converges to Q if the output levels of the Q_n's converge to the output levels of Q. Notice that this definition allows the limiting quantizer Q to have fewer levels than the quantizers Q_n.

Let F_n and F be k-dimensional probability distribution functions. Recall that the sequence $\{F_n\}$ is said to converge weakly to F (written $F_n \xrightarrow{w} F$) if $F_n(x) \to F(x)$ at every continuity point x of F.

A function C: $\mathbb{R}^k \to \mathbb{R}$ is said to be lower semi-continuous if

$$\lim_{\delta \to 0} \inf_{0<\|y-x\|<\delta} C(y) \geq C(x) \text{ all } x.$$

An equivalent condition is that for any $\{x_n\}$ and x, $x_n \to x$ implies

$$\lim\inf C(x_n) \geq C(x).$$

A continuous function g: $\mathbb{R}^k \to \mathbb{R}$ is said to be uniformly integrable with respect to the distributions $\{F_n\}$ if $\int |g(x)| dF_n(x) < \infty$ for all n and

$$\lim_{a \to \infty} \sup_n \int_{\|x\|>a} |g(x)| dF_n(x) = 0.$$

II. Existence of Optimal Quantizers

In this section we review some existence theorems for globally optimal quantizers, and then prove some stronger results. The first two theorems below are proved in [7,8]. The present statement of Theorem 1 is a slight but straightforward extension of that proven in [8].

Theorem 1. Suppose that the cost function $C_o(t)$ is non-negative, non-decreasing for $t \geq 0$, and lower semi-continuous. Then for any distribution F the distortion $D(y_1,y_2,\ldots,y_N)$ is a lower semi-continuous function of the output levels. If $C_o(t)$ is continuous, then $D(y_1,y_2,\ldots,y_N)$ is continuous. Moreover, if F is continuous, then $D(y_1,y_2,\ldots,y_N)$ is continuous, even if $C_o(t)$ is not required to be semi-continuous.

Theorem 2. Let F be a given k-dimensional distribution function. If $C_o(t)$ is non-negative, non-decreasing for $t \geq 0$ and lower semi-continuous, then a globally optimum N-level quantizer exists for any positive integer N.

This theorem gives a sufficient condition on the form of the cost function $C_o(t)$ so that (1) may achieve its minimum possible value. The distribution F is completely general, and all of the assumptions are directed towards the cost function. The key assumption here is lower semi-continuity; the other two conditions merely serve to define the intuitive notion of a difference-based cost function. These conditions may not be weakened without adding other qualifications, as the following example shows. Let the cost function be defined as

$$C_o(t) = \begin{cases} 0 & 0 \leq |t| < 2 \\ 1 & 2 \leq |t| \end{cases} . \tag{2}$$

Now let X take on a discrete distribution such that

$$P(X=2) = P(X=-2) = 1/4$$

$$P(X=2-1/k) = P(X=-2+1/k) = 1/2^{k+2} \quad k=1,2,\ldots \tag{3}$$

It is easy to see that with N=1 level, the distortion (1) cannot be minimized. In fact, any reasonable numerical technique for optimizing (1) will generate a sequence of quantizers with improving distortions, but the limiting quantizer exhibits worst-case performance. Note that we are concerned here with global minima, and not merely local minima or stationary points, of which this example has infinitely many.

This example can be generalized. Suppose that $C_o(t)$ is non-negative, non-decreasing for $t \geq 0$, and is not lower semi-continuous at $t_o > 0$. Then the discontinuity (of the first kind) at t_o must have the same character as that in (2), namely, $C_o(t_o-) < C_o(t_o)$, and it will be possible to construct a distribution with atoms at $\pm t_o$, such that the mean distortion (2) cannot have a minimum. Thus if we wish to avoid this situation, we must require $C_o(t)$ to be lower semi-continuous. On the other hand, it has been shown in Theorem 2 that lower semi-continuity is sufficient to guarantee the existence of a global minimum. Therefore we have the following theorem.

Theorem 3. Suppose that $C_o(t)$ is non-negative and non-decreasing for $t \geq 0$. Then a necessary and sufficient condition that, for any N and any distribution function, a minimum distortion N-level quantizer exists, is that $C_o(t)$ be lower semi-continuous.

In some situations there might be some uncertainty about the distribution F being quantized. This theorem says that if we choose a lower semi-continuous cost function, then no matter what distribution might be involved, we can be sure that the distortion achieves a minimum. Moreover, this assurance vanishes if the function is not lower semi-continuous. All of the popular cost functions, such as the r^{th}-power cost function, satisfy the requirements of Theorem 3.

Theorem 3 characterizes a class of cost functions so that (2) can be minimized universally.

The converse question is also of some interest. That is, we wish to characterize a class of distributions for which the same property holds. A partial answer may be found in Theorem 1. For if the distribution function F is continuous, then $D(y_1, y_2, \ldots, y_N)$ will be a continuous function of the output levels. Using the same methods in [8] that were used to prove Theorem 2, it can be shown that the mean distortion is minimized by some N-level quantizer. This gives the following sufficiency condition.

Theorem 4. Suppose that the distribution F of the random vector being quantized is continuous. Then for any cost function $C_o(t)$ that is non-negative and non-decreasing for $t \geq 0$, a globally minimum distortion N-level quantizer exists for every N.

For example, the well-known Cantor distribution [11] is continuous, although singular with respect to Lebesgue measure. Nevertheless, any reasonable cost function yields a distortion function that can be minimized. Theorem 4 is satisfying because in many practical situations the random vector being quantized has an absolutely continuous distribution, and so a large class of difference-based cost functions is available.

III. Uniqueness and Symmetry of Quantizers

The results of the previous section are valid for both vector and scalar quantizers. In this section, we shall be mostly concerned with scalar quantizers. Through examples, we shall point out some interesting aspects about symmetry and uniqueness of optimal quantizers.

Many popular univariate density functions possess even symmetry, e.g., the zero mean Gaussian density, the Laplace density and the generalized Gaussian densities [12]. When this symmetry is present, it is a common practice to look for an optimum symmetric scalar quantizer, i.e., one whose output levels are symmetrically disposed about the origin [4,5]. Since in this case only half of the output levels have to be computed, considerable computational savings are possible, especially when the number of levels N is large. Unfortunately, it is not true that the optimal quantizer for a symmetric density function is itself symmetric. So this procedure can result in a sub-optimal design. Consider the density given by

$$f(x) = \begin{cases} -|x|/3 + 5/12 & |x| < 1 \\ 7 - |x|/72 & 1 \leq |x| < 7 \\ 0 & 7 \leq |x|. \end{cases} \quad (4)$$

This density has two minimum mean squared two-level quantizers Q(x) and Q(-x) given by

$$Q(x) = \begin{cases} -1 & x < 1 \\ 3 & x \geq 1. \end{cases}$$

These have mean squared errors D = 2.61. By comparison the best symmetric two-level quantizer for f(x) has output levels at $\pm$ 61/36 and mean squared error D = 2.74. Similar examples can be constructed for any number of levels N and for other cost functions. This example makes it clear that we cannot in general hope to simplify the problem of constructing optimal quantizers by appealing to symmetry. In applications, usually the probability density to be quantized is not precisely known. Typically, the greatest uncertainty is associated with the tails of the density, while the central part may be well estimated. If the nominal density f_o is symmetric, the actual distribution may be any member of the mixture class of symmetric densities

$$\{(1-\varepsilon)f_o + \varepsilon h:\ h(x) \text{ is a symmetric univariate density}\}$$

for some positive ε. We show that with an r^{th}-power cost function, there is some member of this mixture class of symmetric densities which has a non-symmetric optimal quantizer. First we state a theorem from [9] which will be used in the subsequent proof. This theorem holds for vector quantizers.

Theorem 5. Assume that $C_o(t)$ is a non-constant function which is non-negative, non-decreasing for $t \geq 0$ and continuous. Suppose that $F_n \xrightarrow{w} F$, and that Q_n is an optimal N-level quantizer for F_n. If $C(x-y)$ is uniformly integrable with respect to $\{F_n\}$ for each $y \in \mathbb{R}^k$, then every limiting quantizer Q of the sequence $\{Q_n\}$ is optimal for F.

Now we can state and prove a theorem for scalar quantizers.

Theorem 6. Let the cost function be given by $C_o(t) = t^r$ for some positive r. Let N be a positive even integer and let f_o be a given symmetric density function with finite r^{th} absolute moment. For any given ε, $0 < \varepsilon < 1$, there exists a symmetric density function h(x) such that for the mixture density $(1-\varepsilon)f_o + \varepsilon h$, no optimal N-level scalar quantizer is symmetric.

Proof: For simplicity we take N=2. The proof generalizes easily to higher even N. Define

$$h_n(x) = \begin{cases} (|x| - n+1)/2 & n-1 \leq |x| < n \\ -(|x| - n-1)/2 & n \leq |x| < n+1 \\ 0 & \text{elsewhere.} \end{cases} \quad (5)$$

This probability density function consists of two triangular masses centered at $\pm n$. Now form the mixture density

$$f_n(x) = (1-\varepsilon)nf_o(nx) + \varepsilon nh_n(nx).$$

As n becomes large, the contribution of f_o becomes probability mass around the origin, and that of h_n becomes two probability masses equally distributed around ± 1. In other words, the distribution represented by f_n converges weakly to a distribution with the cdf:

$$F(x) = \begin{cases} 0 & x < -1 \\ \varepsilon/2 & -1 \leq x < 0 \\ 1-\varepsilon/2 & 0 \leq x < 1 \\ 1 & 1 \leq x \end{cases}.$$

It is a simple matter to verify that $C(x-y) = |x-y|^r$ is uniformly integrable with respect to $\{F_n\}$, where F_n is the distribution function of f_n. Thus Theorem 5 can be applied. Taking $\{Q_n\}$ to be an arbitrarily chosen sequence of optimum 2-level quantizers for $\{F_n\}$, one sees that every convergent subsequence of $\{Q_n\}$ must approach an optimum configuration of levels for F. It is easily seen that there are exactly two optimal 2-level quantizers for F. For example, with a mean square error criterion (r=2) the optimal sets of output levels are $\{-\varepsilon/(2-\varepsilon), 1\}$ and $\{-1, \varepsilon/(2-\varepsilon)\}$, which are clearly non-symmetric, especially for small ε.

Since the limit quantizers are non-symmetric, there must be infinitely many densities f_n having non-symmetric optimum quantizers Q_n. The optimal quantizer for $(1-\varepsilon)f_o(x) + \varepsilon h(x)$ is obtained by scaling the output levels of Q_n by the factor n. Thus there are infinitely many h_n satisfying the requirements of the theorem.

The contaminating density $h_n(x)$ need not have the triangular shape in (5). It could just as easily have been made infinitely differentiable. For example, two Gaussian bell-shaped curves centered at $\pm n$ would have given the same result. The phenomenon described in the theorem has been noted previously [7]. However, the present version is more general than the earlier one.

Having seen these examples, we might be interested in finding some condition so that a symmetric probability density will have a symmetric optimum quantizer. To the authors' knowledge, the only condition which guarantees this is uniqueness of the optimum quantizer. For if a density has a single locally optimum quantizer, then that quantizer must be symmetric. Fleischer [6] derived a simple criterion that guarantees uniqueness of the stationary point in the case of mean square distortion.

Theorem 7. Suppose that a twice-differentiable positive univariate density function f with finite second moment satisfies the inequality

$$\frac{d^2}{dx^2} \ln f(x) < 0 \tag{6}$$

for all x. Then the mean squared error $D(y_1, y_2, \ldots, y_N)$ of an N-level quantizer (N arbitrary) has a unique stationary point, and this stationary point is a global minimum.

Many densities of practical interest, including the Gaussian density, satisfy (6). However, one that does not meet the criterion is the Laplace density [12]

$$f(x) = 1/2\, e^{-|x|} .$$

By taking the Laplace density as the uniform limit of a sequence of densities ($\alpha \to 0$)

$$f_\alpha(x) = 1/2\, e^{-|x|}\left[1 - \alpha e^{-\frac{1-\alpha}{\alpha}|x|}\right] \tag{7}$$

Fleischer showed that it also possesses a unique minimum mean square distortion quantizer. One might then wonder about the strictness of Fleischer's condition (6). Can it, for example, be weakened to hold almost everywhere on the real line? The answer is no. For example, the density given in (4) satisfies Fleischer's condition everywhere in [-7,7] that the second derivative exists, yet it has three different stationary points (locally optimum quantizers). By taking a mixture $(1-\varepsilon)f(x) + \varepsilon e^{-x^2/2}/\sqrt{2\pi}$, where f(x) is given in (4), we can get a positive density which satisfies (6) almost everywhere; but again, this mixture density has three distinct locally optimum quantizers. Clearly, some care must be exercised in extrapolating Theorem 7 beyond its stated conditions. The following examples may be illuminating.

Let the cost function be $C_o(t) = |t|$. It is well known that for N=1, the optimum output level is a median. Consider the sequence of densities

$$f_\alpha(x) = \begin{cases} \alpha & 0 \le |x| < 1 \\ 1/2 - \alpha & 1 \le |x| < 2 \\ 0 & \text{elsewhere.} \end{cases} \tag{8}$$

Each density has a unique median and therefore the unique minimum mean absolute deviation 1-level quantizer is $Q(x) \equiv 0$. As $\alpha \to 0$ the uniform limit of the sequence is the density

$$f(x) = \begin{cases} 1/2 & 1 \le |x| < 2 \\ 0 & \text{elsewhere.} \end{cases}$$

It is clear that this density has infinitely many medians, and so the optimum quantizer cannot be unique.

A second example illustrates another difficulty. Let the cost function be $C_o(t) = t^2$, giving the mean squared error criterion. In this case the optimum 1-level quantizer level is the mean. Consider the sequence of densities

$$f_n(x) = (1-\tfrac{1}{n})\left(\frac{1}{\sqrt{2\pi}}\right)e^{-x^2/2} + \frac{1}{n}\left(\frac{1}{\sqrt{2\pi}}\right)e^{-(x-n)^2/2}. \tag{9}$$

As $n \to \infty$ the sequence converges uniformly to the standard Gaussian density

$$f(x) = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} .$$

In fact, $|f_n(x) - f(x)| < 1/n$ uniformly. Nevertheless each $f_n(x)$ has a mean (optimum output level) of 1, whereas it is well known that the optimum 1-level Gaussian quantizer has the output level at 0. So the limiting quantizer is not even locally optimum.

These examples illustrate that uniform convergence of the sequence of densities involved is not enough to assure that the limiting quantizer is either optimal or unique. Some other ingredients are necessary, and these are given in the following theorem and corollary. This theorem is stated for multivariate distributions, but the application to univariate densities is straightforward.

Theorem 8. Assume that $C_o(t)$ is non-negative, non-decreasing for $t \geq 0$, and continuous. Let F be a probability distribution function which has isolated locally optimum quantizers Q_j. Suppose that $\{F_n\}$ is any sequence of distributions converging weakly to F, such that $C(x-y)$ is uniformly integrable with respect to $\{F_n\}$. Then for each Q_j there exists a sequence of quantizers $\{Q_n^j\}$ so that Q_n^j is a locally optimum quantizer for F_n, and $Q_n^j \to Q_j$ as $n \to \infty$.

Proof: Due to the assumption of nearest neighbor assignment, we can represent Q_j by a product vector $\tilde{y}_j = \{y_1, y_2, \ldots, y_N\}$, where the output levels $y_i \in \mathbb{R}^k$ are indexed in order of increasing norm. Since the Q_j are isolated locally optimum quantizers, there exists, for each j, an ε_{oj} and a closed ball

$$\bar{N}_{\varepsilon_{oj}}(Q_j) = \{\tilde{y} \in \mathbb{R}^{kN} : \|\tilde{y}-\tilde{y}_j\| \le \varepsilon_{oj}\}$$

in which $\tilde{y}_j$ is the only minimum of $D(Q,F)$. Henceforth j will be a fixed integer and we may let $\varepsilon_o = \varepsilon_{oj}$. We will prove the following proposition:

For all $0 < \varepsilon < \varepsilon_o$, there exists an integer n_o such that for $n \ge n_o$, each of the $D(Q,F_n)$ have a local minimum point inside the open ball

$$N_\varepsilon(Q_j) = \{\tilde{y} \in \mathbb{R}^{kN} : \|\tilde{y}-\tilde{y}_j\| < \varepsilon\}$$

centered on $\tilde{y}_j$.

The proof is by contradiction. Suppose the proposition were not true for some positive $\varepsilon < \varepsilon_o$. Then we could find a subsequence $\{F_{n_k}\}$ and quantizers Q^k so that Q^k is on the boundary of $N_\varepsilon(Q_j)$

$$D(Q^k, F_{n_k}) < D(Q, F_{n_k}) \text{ for } Q \in N_\varepsilon(Q_j).$$

This is so because the distortion $D(Q,F_{n_k})$ is a continuous function of the output levels (Theorem 1); so that if it has no minimum inside $N_\varepsilon(Q_j)$, then the minimum (corresponding to Q^k) must occur on the boundary. The sequence $\{Q^k\}$ has a limit quantizer Q^∞ on the boundary, and a slight modification of the proof of Theorem 5 yields that Q^∞ is optimal for F among the quantizers restricted to $\bar{N}_{\varepsilon_o}(Q_j)$. But this contradicts the choice of ε_o, which proves the proposition.

By taking a sequence $\varepsilon_m \downarrow 0$, we conclude that every subsequence of $\{F_n\}$ contains a further subsequence F_{n_m}, for which there is a sequence $\{Q^j_{n_m}\}$ so that $Q^j_{n_m} \to Q^j$ and each $Q^j_{n_m}$ is locally optimum for F_{n_m}. Therefore the conclusion of the theorem must be true for the entire sequence $\{F_n\}$.

According to the theorem, if $f_n \to f$ and the density f has m distinct locally optimum quantizers, then every f_n with n sufficiently large has at least m local minima. Now if each of the f_n's has a unique stationary point, it must be true that f has exactly one stationary point. This gives the following corollary.

Corollary 9. Assume that $C_o(t)$ is non-negative, non-decreasing for $t \ge 0$, and continuous. Let f_n, f be density functions such that $f_n \to f$ and $C(x-y)f_n(x)$ is uniformly integrable for all y (with respect to Lebesgue measure). Suppose that each of the f_n's has a unique optimum quantizer. Then either f has a unique optimum quantizer, or it has several minima which are not separated.

Going back to the previous examples, we can see that in (8) the local minima (medians) of the limiting density are not isolated. In (9), the uniform integrability assumption is not satisfied. Needless to say, both of these requirements are met by Fleischer's construction (7).

IV. Conclusion

We have studied the questions of existence, uniqueness and symmetry of optimal quantizers with difference-based distortion measures. Two theorems are proved which guarantee existence of an optimal quantizer: the first a necessary and sufficient condition on the form of the cost function and the second a sufficient condition on the distribution. Then we investigated a way to prove uniqueness of a local optimum by considering sequences of densities. Several examples are presented to clarify the discussion.

As a final comment, we note that the existence results in Theorems 3 and 4 can be straightforwardly extended to scalar uniform step size quantizers, or to scalar symmetric quantizers.

Acknowledgement

This research was supported by the Air Force Office of Scientific Research under Grant AFOSR-81-0047.

REFERENCES

[1] J. Max, "Quantizing for Minimum Distortion," IRE Trans. Inform. Theory, IT-6 (1960), pp. 7-12.

[2] Y. Linde, A. Buzo and R.M. Gray, "An Algorithm for Vector Quantizer Design," IEEE Trans. Comm., COM-28 (1980), pp. 84-95.

[3] J.D. Bruce, "On the Optimum Quantization of Stationary Signals," IEEE International Convention Record, Pt. I, pp. 118-124, 1964.

[4] M.D. Paez and T.H. Glisson, "Minimum Mean-Squared Error Quantization in Speech PCM and DPCM Systems," IEEE Trans. Comm., COM-20 (1972), pp. 225-230.

[5] P.F. Panter and W. Dite, "Quantization Distortion in Pulse-Count Modulation with Non-uniform Spacing of Levels," Proc. IRE, vol. 39 (1951), pp. 44-48.

[6] P. Fleischer, "Sufficient Conditions for Achieving Minimum Distortion in a Quantizer," IEEE International Convention Record, Pt. I, pp. 104-111, 1964.

[7] E.F. Abaya and G.L. Wise, "Some Notes on Optimal Quantization," Proc. Int. Conf. on Comm., vol. 2, pp. 30.7.1-30.7.5, June 14-17, 1981.

[8] E.F. Abaya and G.L. Wise, "On the Existence of Optimal Quantizers," to appear in IEEE Transactions on Information Theory, November, 1982.

[9] E.F. Abaya and G.L. Wise, "Convergence of Vector Quantizers with Application to the Design of Optimal Quantizers," Proceedings 19th Allerton Conference on Communications, Control, and Computing, pp. 79-88, Sept. 30-Oct. 2, 1981.

[10] S.P. Lloyd, "Least Squares Quantization in PCM," Bell Telephone Laboratories Paper, Murray Hill, N.J., 1957.

[11] R.G. Laha and V.K. Rohatgi, Probability Theory, John Wiley and Sons, New York, 1979.

[12] J.H. Miller and J.B. Thomas, "Detectors for Discrete-Time Signals in Non-Gaussian Noise," IEEE Trans. Inform. Theory, IT-18 (1972), pp. 241-250.

ATE

LMED

-8