Fast Index Assignment for balanced N-Description Scalar Quantization

Ivana Radulovic and Pascal Frossard

Signal Processing Institute, EPFL

Lausanne 1015, Switzerland

{ivana.radulovic, pascal.frossard}@epfl.ch

Fax: +41 21 693 7600, Phone: +41 21 693 2601

Abstract

In this paper, we address the design of *any* number of balanced descriptions using multiple description scalar quantization(MDSQ), based on fast index assignment. The proposed scheme has the advantages of low complexity, and the possibility of being extended easily to any number of descriptions. Unlike existing schemes, it can produce balanced descriptions even at low rate, at the price however of a slightly higher distortion. The behavior of the proposed index assignment at high rate is in the same time similar to state-of-the-art schemes. The proposed scheme offers the possibility to adapt to loss probability, and rate constraints, in playing with both the number of descriptions, and the rate of each of them, to minimize the average distortion.

I. INTRODUCTION

In general, the goal of multiple description coding (MDC) is the generation of two or more descriptions of a source, such that each of them gives acceptable reconstruction quality when received alone, and the representation of source improves when the number of received descriptions increases. Under a generic form, each description n has a rate R_n, and a side distortion D_n. In this paper, we are addressing a specific case where all the descriptions are balanced, i.e., where all the rates R_n are equal. Moreover, in the case of a uniform source, all the side distortions are equal, and the distortion after receiving any k among the N distortions are also equivalent.

Multiple description scalar quantization (MDSQ) has given the first practical solution for generating multiple descriptions. The design of an MDSQ system generally follows two steps, a scalar quantization and an index assignment method, that maps the quantized value to a N-tuple of quantization indices $(i_1, i_2, ..., i_N)$. Each index is then sent over a different channel: if all the indexes are correctly received, the signal will be reconstructed with the highest quality that is given by the scalar quantization step. If any channel fails, the decoder still reconstructs a version of the signal, with a lower quality.

In this paper, we propose two simple index assignment schemes able to generate *any* number of *balanced* descriptions. Except for unequal error protection (UEP) based schemes, the common MDC schemes are generally limited to the generation of three descriptions. The proposed design easily extends to any number of descriptions, while keeping a very low complexity. We show that, in the case of uniformly distributed sources,

we can achieve balanced descriptions both at low and high rates, which is not possible with state-of-the-art MDSQ schemes. The proposed scheme is later extended to the case of Gaussian sources. The proposed schemes are shown to perform similarly to state-of-the-art encoders at high rate. The side distortion is however slightly penalized by the condition of balancing distortions.

Finally, the fact that we can generate any number of descriptions is advantageously used to optimize the average distortion for given loss probabilities and rate constraints. It provides the flexibility to play with both the coding rates, and the number of descriptions, without being penalized by the so-called cliff-effect that is observed in UEP-based solutions.

The paper is organized as follows: Section II briefly overviews the related work. Section III describes our approach for the generation of n balanced descriptions. Section IV considers the special case of three descriptions, which interestingly presents particular possibilities of extensions, different than the general scheme. Comparisons of our scheme with state-of-the-art methods are given in Section V. Section VI considers the problem of minimizing the average distortion, under given loss probability and rate constraints. Finally, conclusions are gievn in Section VII.

II. RELATED WORK

The first approach for the generation of two balanced descriptions used the staggered index assignment principle: two side quantizers are half of the quantization step size offset to each other so that receiving both descriptions results in four times smaller distortion. This scheme has a high degree of redundancy, since for B-bit resolution we need $2(B-1)$ bits. Therefore, it is a good solution only in the case when we want to minimize the side distortion.

In order to reduce the redundancy between descriptions, the authors in [7] [8] proposed using disjoint quantization bins. Moreover, the whole family of diagonal index assignment matrices was proposed. By varying the number of diagonals in such matrices, one can control the tradeoff between the side and central distortions. Although the presented results were remarkable, the design of more than two descriptions was not considered. Extending this method to the design of n descriptions would require the search for the solution in the hypercube of dimension n, which is not a trivial problem. This family of index assignments was successfully applied in wavelet-based image coders, [6].

In parallel to scalar quantization, other ideas were proposed for the generation of more than two balanced descriptions. Most of them were based on unequal error protection (UEP) principles. In [1], the authors use Reed-Solomon (n,k) codes to make equally important descriptions from the output of a progressive coder. In [2], they propose adding the controlled amount of redundancy to the same progressive coder. This is achieved by spreading the information about each wavelet tree in many descriptions, which guarantees recovery of the most important information. Schemes based on UEP assume that at least k out of n descriptions are received and they don't reconstruct the signal if fewer than k descriptions are received. Moreover, receiving more than k descriptions might not bring any improvement in the reconstructed quality. This is called the cliff effect. Besides, the protection level of these schemes ($\frac{k}{n}$) depends on the state of the channel, which may change very rapidly during the transmission. In order to soften the cliff effect, the authors in [5] propose using (n,k) source-channel erasure codes for the generation of n descriptions. They also derive the complete rate region for their scheme. However, their study is still based on the assumption that at least k descriptions are received.

Recently, attention has been put again to the problem of generating more than two descriptions with scalar quantization. The staggered index assignment, combined with UEP, was proposed for the design of three balanced descriptions in [5] and as the first stage of the design in [3]. However, these schemes fail to give balanced descriptions at low rates. Moreover, in [3] they stress that their method is not proper for the design of more than three descriptions because of only one degree of freedom in the index assignment matrix. In [4], the author proposed the encoding procedure in a multistage fashion, where each stage doubles the number of descriptions using the method proposed in [7]. However, his scheme allows only the number of descriptions that is a power of two.

In this paper we tackle the problem of generating balanced descriptions based on scalar quantization, and we give a simple and easily extendable scheme that shows how to generate N balanced descriptions.

III. Index Assignment for Balanced MDSQ
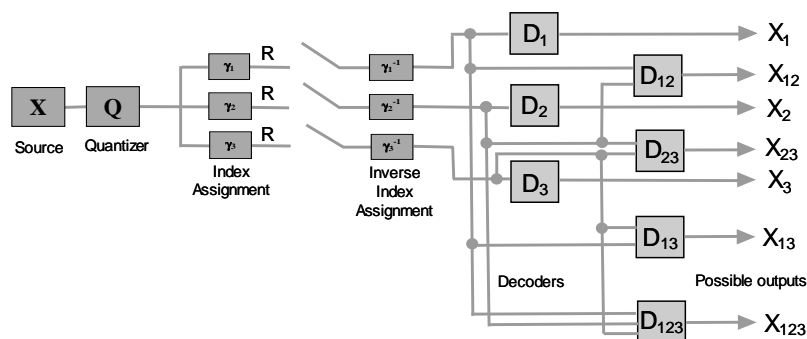
A. Preliminaries


Fig. 1.   Three Description Coding Scheme.

In this section, we consider the problem of index assignment, for the generation of N balanced descriptions, based on scalar quantization. The case of three descriptions can be represented as in Figure 1. Each of the N description has a rate R_n, and the total rate R therefore becomes $R = \sum_{n=1}^{N} R_n$. The distortion of a single description is called the side distortion, D_n. The reception of all the descriptions induces a central distortion D, and the reception of only a subset $\mathcal{A}$ of the descriptions is denoted partial distortion, $D_{\mathcal{A}}$. The reconstruction points are chosen as the centroid of the interval that is common to the received description bins. The distortion is measured by the mean square error, MSE, relative to the input signal.

We concentrate in this section on two different types of sources: a uniform source on the interval $[0, 1]$ and the Gaussian source. In the uniform case, we assume that the MDSQ scheme first performs a uniform quantization of the input signal, with a step-size δ, without loss of generality. Then, index assignment attributes indexes to each of the uniform quantization intervals, that can be different for each of the descriptions. Index assignment becomes the key of the MDSQ scheme, and directly drives the distortion when descriptions are lost. In the lossless case, a proper index assignment should result
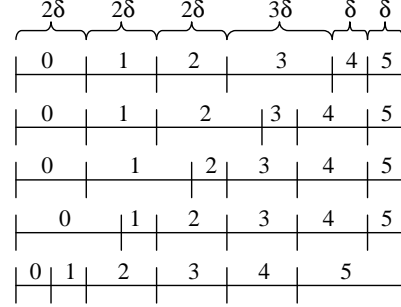
in a uniform quantization of the source, with step-size δ, when all N descriptions are combined together.

The remainder of the section now describes two fast index assignment schemes for the generation of N balanced descriptions. The methods are further extended to Gaussian sources, when all the decision levels are adapted to the different distribution.

B. Balanced index assignment for N-description: uniform distribution
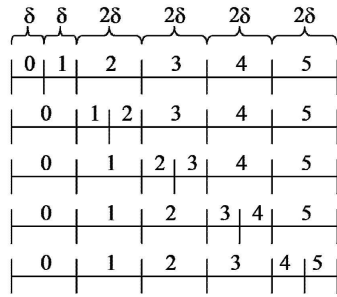
(a) Step Split algorithm.   (b) Merge and Split algorithm.

Fig. 2.   Index assignment strategies for $N = 5$.

1) Step split scheme: To generate N balanced descriptions, we start from N uniform quantizers with step-size 2δ. One way of generating them consists in dividing i-th bin of i-th quantizer into two bins of size δ. Such a scheme is represented in Figure 2 (a) for the case of $N = 5$, and it can be seen that the reception of all the descriptions guarantees an uniform scalar quantization with step-size δ.
All the descriptions obviously have the same rate R_n or entropy, since they all have the same set of quantization bins Δ_i, $R_n = -\sum_i \Delta_i log \Delta_i$. It can be further verified that the combination of *any* k out of N descriptions will result in $2k$ bins of size δ and $(N-k)$ bins of size 2δ. This will additionally guarantee balanced partial distortions.

In order to increase the bit rate, the (basic) structure explained above can be simply periodically repeated. Finally, if the number of descriptions is N and the number of repetitions is m, we will have $\delta = \frac{1}{2Nm}$ and the description rate, respectively the partial distortion of receiving k out of N descriptions, are given by :

$$R_n = log(2Nm) - \frac{N-1}{N} \tag{1}$$

$$D_{12...k} = \frac{4N-3k}{48m^2 N^3} = \frac{4N-3k}{12N2^{\frac{2(N-1)}{N}}} 2^{-2R}, k \geq 1. \tag{2}$$

2) Merge and split scheme: Another way to generate N balanced descriptions is based on a similar idea, but instead of dividing one bin of size 2δ into two bins of size δ, as proposed above, the index assignment merges two bins of size 2δ and then splits the new bin into one bin of size δ and another of size 3δ. The *Merge and split* algorithm results in a structure that has $N-2$ consecutive intervals of size 2δ, one of size 3δ, and one of size δ. Cyclic permutation of such a structure, by steps of 2δ, generates the set of N

descriptions, where one column of size δ has finally to be added to compensate for edge effects. The resulting index assignment is represented in Figure 2(b) for the case of five descriptions.

It can be shown that the combination of any k out of N descriptions results in $(2k+1)$ bins of size δ and $(N-k)$ bins of size 2δ. Different bit rates can again be obtained by simple repetitions of the building block defined hereabove. Thus, for N descriptions and m repetitions of the basic structure, the Merge and split scheme yields $\delta = \frac{1}{m(2N+1)}$ and the description rate, side and partial distortions of receiving any k out of N descriptions, are given by :

$$R_n = log(m(2N+1)) - \frac{2(N-2) + 3log3}{2N+1} \tag{3}$$

$$D_n = \frac{8N+13}{12m^2(2N+1)^3} = \frac{8N+13}{12(2N+1)2^{\frac{4(N-2)+6log3}{2N+1}}} 2^{-2R_n} \tag{4}$$

$$D_{12...k} = \frac{8N-6k+1}{12m^2(2N+1)^3} = \frac{8N-6k+1}{12(2N+1)2^{\frac{4(N-2)+6log3}{2N+1}}} 2^{-2R_n}, k \geq 2. \tag{5}$$

It can be seen that there is an explicit relation between the rate and distortions for our schemes, even at low rates. Besides, we see that all the distortions have the same decay rate (2^{-2R_n}) and that they decay linearly as a function of the number of received descriptions, k.

C. Balanced index assignment for N descriptions: Gaussian distribution

The coding of Gaussian sources is derived from the method described hereabove, by mapping the description quantizers to the novel distribution of the source. Denote by $Q_{U,N}$ the set of N quantizers used for the uniform source and by $Q_{G,N}$ the same set for the Gaussian source. The lower and upper decision levels of each bin in $Q_{U,N}$ is transformed to a new set of lower and upper decision levels in $Q_{G,N}$, that will adapt to he Gaussian signal. If the Gaussian source has zero mean and unit variance, the transform is given simply by:

$$g = \frac{1 + erf(\frac{u}{\sqrt{2}})}{2}, \tag{6}$$

where u corresponds to levels in $Q_{U,N}$, and g corresponds to new levels in $Q_{G,N}$. Note that such a transformation guarantees that the rates of the descriptions stay balanced for the Gaussian source. However, the side and partial distortions are not balanced any more at low rate. The reason is that trying to keep the linear metrics balanced (i.e. rates) for the Gaussian distribution, does not allow for keeping the square metrics (i.e., MSE distortion) balanced, at the same time. Finally, a similar transform could also be applied to other source distributions.
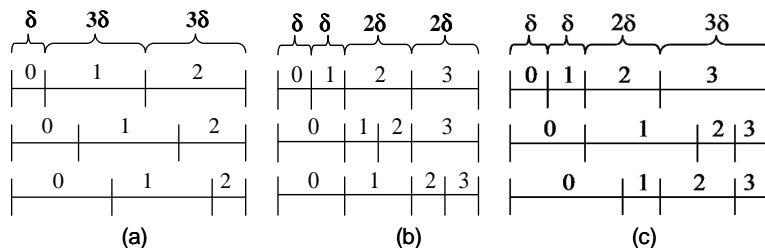
Fig. 3. Schemes for 3 balanced descriptions:(a) existing scheme, (b) scheme ***step split***,(c) scheme ***merge and split***.

IV. SPECIAL CASE: $N = 3$

This section now analyzes the special case of 3 descriptions, and compares the index assignment methods proposed before (see Figure 3 (b) and (c)) to the scheme based on staggered index assignment, that is the first stage of the design in [3] (see Figure 3 (a)).

Consider first the staggered index assignment, represented in Figure 3(a). We see that the quantization bins in different descriptions are offset one third of the quantization step size δ to each other. Receiving any two descriptions will give three bins of size δ and two bins of size 2δ. If the third description is also received, it will solve the ambiguity about bigger intersection bins that exists when only two descriptions are received. The first and the third description have two bins of size 3δ and one bin of size δ, while the second one has two bins of size 2δ and one bin of size 3δ. Therefore, these descriptions do not have the same rates and, moreover, they will result in unbalanced side distortion, i.e., $D_1 = D_3 \neq D_2$. On the contrary, our schemes shift the whole quantization bins instead of just shifting each bin by one third of the quantization step size. All quantizers have therefore always the same set of quantization bins, and rate and distortion stay fully balanced.

Note that in special case of 3 descriptions, there is another possibility to extend the bit rate for the *merge and split* scheme. Instead of repeating the basic structure m times, we can extend our scheme only with intervals of size 3δ, in the very same manner like in the existing scheme. If the number of bins of size 3δ is m, then $\delta = \frac{1}{3m+4}$. Now any two descriptions will have $m + 4$ intersection bins of size δ and m intersection bins of size 2δ. We can calculate the rates and the distortions for this special case, that become:

$$R_n = log(3m + 4) - \frac{2 + 3mlog3}{3m + 4} \tag{7}$$

$$D_n = \frac{27m + 10}{12(3m + 4)^3} \tag{8}$$

$$D_{12} = \frac{9m + 4}{12(3m + 4)^3} \tag{9}$$

$$D_{123} = \frac{1}{12(3m + 4)^2} \tag{10}$$

For this scheme, that is called $MS2$, there is unfortunately no explicit relation between rate and distortions at low rates. However, under the high rate assumption, we will have

$R_n \rightarrow -log(3\delta)$ and:

$$D_n \approx 0.083 \cdot 2^{-2R} \qquad (11)$$

$$D_{12} \approx 0.028 \cdot 2^{-2R} \qquad (12)$$

$$D_{123} \approx 0.009 \cdot 2^{-2R} \qquad (13)$$

As we will show in the next section, this scheme will perform best out of all our proposed solutions and outperforms existing MDSQ schemes.

V. RESULTS

In this section, we compare first the rate-distortion performance of our schemes for the general case of $N \geq 2$ descriptions. Recall that, besides UEP-based methods, there is no other scheme for generation of any number of balanced descriptions.

Let m_1 and m_2 be the numbers of repetitions in schemes *step split* and *merge and split* respectively. For the same rate R, we have the following relation between m_1 and m_2:

$$\frac{m_2}{m_1} = \frac{2N}{2N+1}2^{\frac{2(N-2)+3log3}{2N+1} - \frac{N-1}{N}} > 1. \qquad (14)$$

Now, let us compare $D_{12...k}, k \geq 2$ for schemes *step split* and *merge and split* :

$$\frac{D_{12...k}^{merge\ and\ split}}{D_{12...k}^{step\ split}} = \frac{48m_1^2N^3(8N-6k+1)}{12m_2^2(4N-3k)(2N+1)^3} < (1+\frac{1}{2N})\frac{N^3}{(N+\frac{1}{2})^3}\frac{m_1^2}{m_2^2} < (\frac{2N}{2N+1})^2 < 1. (15)$$

The partial distortion is lower in the *merge and split* scheme. However, in a similar way, it can be shown that the side distortion is bigger for the *merge and split* scheme, i.e., $\frac{D_1^{merge\ and\ split}}{D_1^{step\ split}} > 1$. Since the *merge and split* scheme performs better for all the distortions except the side distortion, it will be used in the remainder of the paper. Note that there is one more way to explain the advantage of the *merge and split* scheme over the other one. Namely, if we define the redundancy as:

$$\rho = NR - R^* = (N-1)log(m(2N+1)) - \frac{N(2(N-2)+3log(3))}{2N+1} \qquad (16)$$

where $R^* = log(m(2N+1))$ is the rate of the single description that causes the same distortion like the central distortion D, we can see that, for the same rate R, the amount of redundancy in scheme *step split* is higher:

$$\rho_{step\ split} - \rho_{merge\ and\ split} = log((2N+1)m_2) - log(2Nm_1) > 0 \qquad (17)$$

Now we consider the special case when the number of descriptions is $N = 3$. We present the results for the partial distortion D_{12} (Figure 4 (a)). It is interesting to notice here that schemes *merge and split* and $MS2$ start on the same point on the RD curve, but scheme $MS2$ will always give better results for higher rates. If we substitute $N = 3$ in (4) and (7), we will see that the scheme $MS2$ will asymptotically perform better since it gives $D_{12} = 0.028 \cdot 2^{-2R}$, while schemes *merge and split* and *step split* give $D_{12} = 0.067 \cdot 2^{-2R}$ and $D_{12} = 0.041 \cdot 2^{-2R}$ respectively. Another feature that makes the scheme $MS2$ superior in the case of three descriptions is the smaller granularity at lower rates.

Next, we compare the performance of the scheme $MS2$ with the one proposed in [5] and as the first stage of the scheme presented in [3]. Recall that the existing schemes *cannot* produce balanced descriptions at low rates, while our scheme produces both balanced rates and distortions for uniform sources. However, at low rates, our scheme performs slightly worse, as can be seen from Figure 4 (b). It can be also derived that the schemes will nevertheless perform the same at high rates.
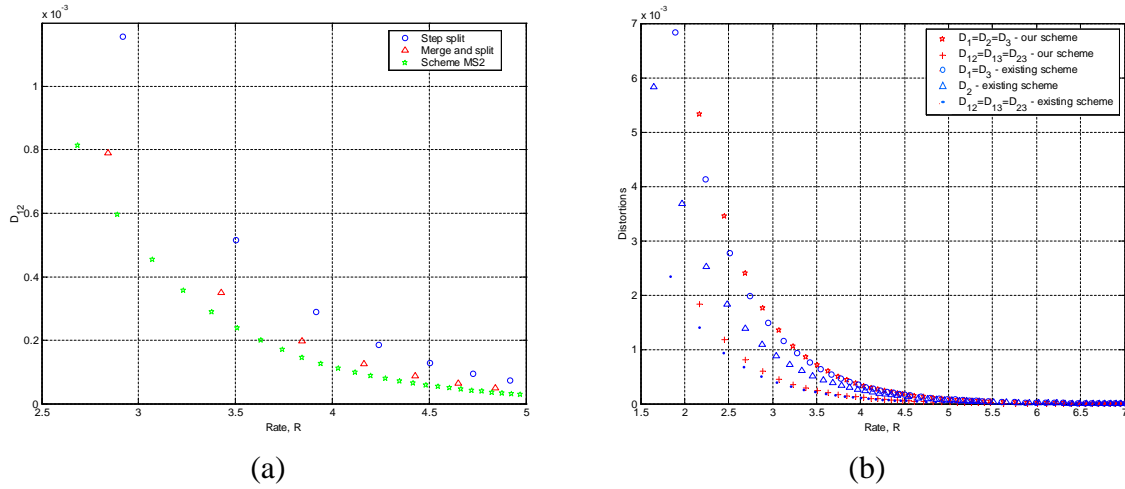

Fig. 4. Comparison of schemes *step split*, *merge and split* and MS2 for three descriptions and distortion D_{12}.
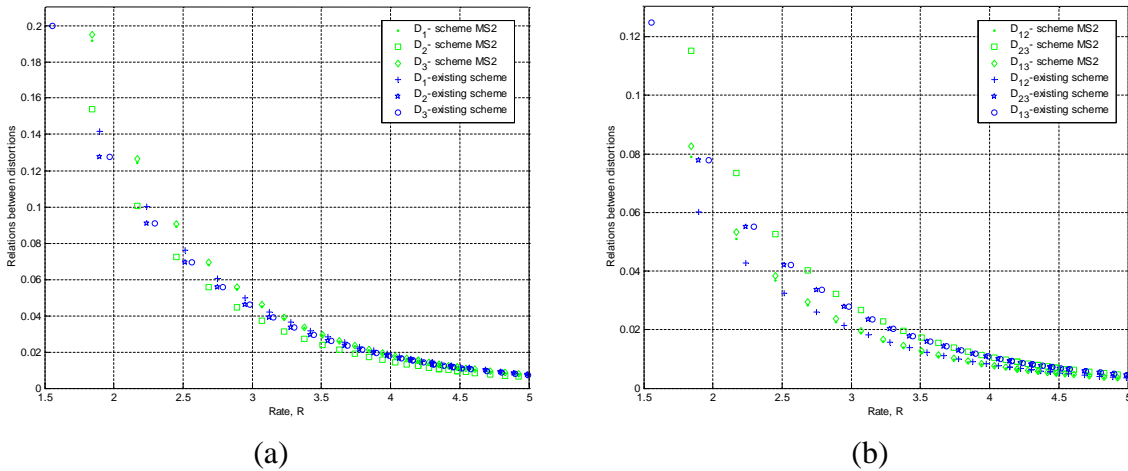

Fig. 5. Comparison of scheme $MS2$ and the existing scheme for the Gaussian distribution and D_i and D_{ij}, $i, j = 1, 2, 3$.

Finally, we compare scheme $MS2$ and existing schemes for the Gaussian distribution function. The relation between the distortions D_i, $i = 1, 2, 3$ for the two schemes is given on Figure 5(a), while the relation between the distortions D_{ij}, $i, j = 1, 2, 3$ is given on Figure 5(b). From the first figure, we see that D_2 performs slightly better in our scheme, D_3 performs slightly worse, while for D_1 there is a crossing point at low rate, after which

our scheme will slightly perform better than the staggered index assignment. At higher rate, both schemes perform the same. The same conclusion can be drawn for distortions D_{ij} from Figure 5(b). It is interesting to notice here that the distortion in the existing scheme based on staggered index assignment is close to balanced, even though they are not balanced in terms of rate. On the other side, $MS2$ descriptions are completely balanced in terms of rate, but tend to give less balanced distortions at low rates.

VI. MINIMIZATION OF THE AVERAGE DISTORTION IN LOSSY SCENARIOS

This section now briefly addresses the problem of minimization of the average distortion, in lossy scenarios. Assume we can send N descriptions over a lossy network, each one over a different channel. Assume also that the probability that each channel will break down is equal to p.

At the receiver, if we receive all descriptions, we will be able to reconstruct the signal with the lowest distortion (D). If we receive any k descriptions out of N, we will reconstruct the input signal with the distortion $(D_{12...k})$. The most severe case will correspond to the case when all descriptions are lost: in that case, the receiver can just guess what was sent and the distortion will be equal to the variance of the source. Since all the descriptions have the same rate and they are balanced, we can write the average distortion in the presence of losses in the following way:

$$D_{av} = \sum_{k=0}^{N-1} \binom{N}{k} p^k (1-p)^{N-k} D_{12...(N-k)} + p^N \sigma^2 \qquad (18)$$

It can be seen that the average distortion in the lossy scenario will depend on the number of descriptions, rates of descriptions and the probability of error.

Now we can formulate the following optimization problem: given the total rate R and the loss probability p, find the number of descriptions and the rate of descriptions that will minimize the average distortion. The solution to this question is given in Figure 6, which shows the minimal achievable average distortion as a function of the p, and in Table 1 which shows the best number of descriptions, repetitions and redundancy for the proposed scenario.

It can be seen that the case when the rate budget is small is much more sensitive to losses of descriptions. This is due to the fact that we can produce less descriptions at lower rates, but also because these descriptions are less redundant. We also see from the Table 1 that, for a given rate R, with the increase of packet loss ratio, the optimal number of descriptions increases, the number of repetitions decreases and the redundancy increases. This is not a surprising result, since for the very lossy network it makes sense to make and send more descriptions and to make them more redundant.

VII. CONCLUSIONS

In this paper, we proposed a very simple and easily extendable scheme for the generation of N-balanced descriptions. As for the case of 3 balanced descriptions and uniform distribution, we see that we can make them balanced at any rate, which is not the case with the existing scheme. However, our scheme gives slightly higher distortions at low rates. Our scheme can easily be adapted to different lossy scenarios: depending on the probability of error and the total budget rate, we can find the number of descriptions and the rate of the descriptions that will minimize average distortion.

	R_{tot}[bits]	$p=10^{-3}$	$p=10^{-2}$	$p=0.1$	$p=0.2$	$p=0.4$	$p=0.6$	$p=0.9$
N		2	2	2	2	2	2	2
m	4	1	1	1	1	1	1	1
ρ		0.42	0.42	0.42	0.42	0.42	0.42	0.42
N		2	2	3	4	5	5	5
m	15	269	69	8	2	1	1	1
ρ		6.53	6.53	8.72	8.62	8.95	8.95	8.95
N		2	3	4	5	6	8	8
m	25	2239	89	16	5	2	1	1
ρ		11.55	15.67	17.62	18.24	17.62	20.73	20.73
N		3	4	5	6	8	11	11
m	40	2878	223	45	15	3	1	1
ρ		25.7	29.02	30.92	32.15	31.82	34.35	34.35

TABLE I

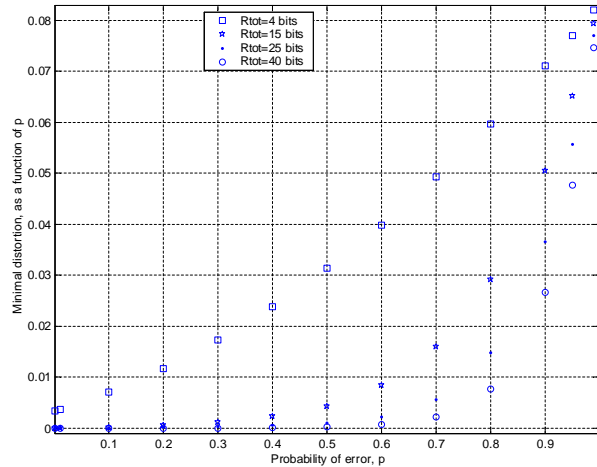OPTIMAL NUMBER OF DESCRIPTIONS AND REPETITIONS AND REDUNDANCY AS A FUNCTION OF p.

Fig. 6. Minimal achievable distortion, as a function of packet loss ratio.

REFERENCES

[1] A.E. Mohr, E.A. Riskin, R.E. Ladner,. Generalized multiple description coding through unequal loss protection. In *Proceedings of the IEEE International Conference on Image Processing, Volume: 1, pp. 411 - 415, Oct. 1999*.

[2] A.C. Miguel, A.E. Mohr, E.A. Riskin,. SPIHT for generalized multiple description coding. In *Proceedings of the IEEE International Conference on Image Processing, Volume:3, pp.842 - 846, Oct. 1999*.

[3] C. Tian and S.S. Hemami,. "Sequential design of multiple description scalar quantizers",. In *Proceedings of the Data Compression Conference, pp.32 - 41, March 2004*.

[4] J. Cardinal. Multistage index assignments for M-description coding. In *Proceedings of the IEEE International Conference on Image Processing, Volume: 3, pp. 249 - 252, Sept. 2003*.

[5] R. Puri, S.S. Pradhan, K. Ramchandran,. n-channel multiple descriptions: theory and constructions. In *Proceedings of the Data Compression Conference, Apr. 2002, pp.262 - 271*.

[6] S.D. Servetto, K. Ramchandran, V.A. Vaishampayan, K. Nahrstedt,. Multiple description wavelet based image coding. *IEEE Trans. Image Processing*, Volume:9(5):pp.813–826, Oct. 2000.

[7] V.A. Vaishampayan,. Design of multiple description scalar quantizers. *IEEE Trans. Inform. Theory* , Vol. 39(3):pp. 821–834, May 1993.

[8] V.A. Vaishampayan and J. Domaszewicz,. Design of entropy-constrained multiple-description scalar quantizers. *IEEE Trans. Inform. Theory*, Vol. 40(1):pp. 245–250, Jan. 1994.