# MULTI LEVEL MULTIPLE DESCRIPTIONS

*Tal A. Beery, and Ram Zamir*

School of Electrical Engineering
Tel Aviv University
talbeery@post.tau.ac.il, zamir@eng.tau.ac.il

**ABSTRACT**

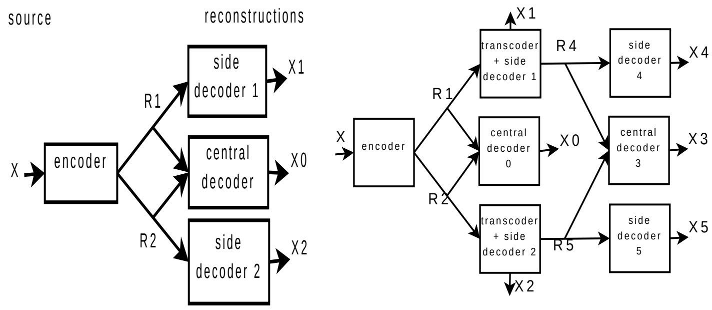
Multiple Description (MD) source coding is a method to overcome unexpected information loss in a diversity system such as the internet, or a wireless network. While classic MD coding handles the situation where the rate in some channels drops to zero temporarily, thus causing unexpected packet-loss, it fails to accommodate more subtle changes in link rate such as rate reduction. In such a case, a classic scheme can not use the link capacity left for information transfer, causing even minor rate reduction to be considered as link failure. In order to accommodate such a frequent situation, we propose a more modular design for transmitting over a diversity system, which can handle unexpected reduction in link's rate, by downgrading the original description into a more coarse description, so it would fit to the new link's rate. The method is analyzed theoretically, and performance results are presented.

***Index Terms—*** Source Coding, Multiple Description Coding, Successive Refinement, lossy packet network, Multiple Description Scalar Quantization.

## 1. INTRODUCTION

All packet based networks suffer from packet-loss. Packet loss occurs when one or more packets of data traveling across a computer network fail to reach their destination. There are numerous reasons why packet loss happens. It could be due to physical change over the network medium such as signal degradation, faulty hardware, or over-saturated network links forcing the router to drop some packets. Notice that the router decision is binary - either pass the packet if there's enough link capacity, or drop it if there's none. If there's only half of the needed capacity, the packet must be dropped. The main measure that networks employ in order to cope with packet loss, is by retransmission. Each packet is numbered, and when a packet is recorded as missing, the receiver asks the sender to retransmit. Although this is a robust method, it has some disadvantages. It's unapplicable over one-way communication scenarios - such as multicast scenarios, and because of the inherent latency in retransmission it's not appropriate for realtime applications.

If the network is a diversity system, i.e. there's more than one way to send information from source to destination, then it may use that diversity to overcome packet-loss. Sending




(a) MD system (b) MLMD system

**Fig. 1**.

packets in different routes over the internet, or transmitting signal over different antennas in a MIMO wireless device, are examples of such. In case that the information can suffer loss and still be intelligible, like audio or video, Multiple Descriptions (MD) is a way to take full advantage of the diversity in the system. A good MD scheme adheres to the following principles:

1. Each description is good by itself. Even if the receiver gets only one description due to packet-loss, the quality of the reconstruction is good enough.

2. If more than one description makes it to the receiver, the outcome has better quality than if it has not.

Clearly, there's an inherent tradeoff between those principles. In order to be better when combined, the descriptions should be as ”different” as possible. But in order to be good by themselves, they should be as ”similar” as possible to the original information.

The inherent tradeoff of the MD problem depicted in fig.1(a) has been examined in information theory perspective, and an inner bound for the achievable rates was found by El Gamal and Cover [1]. An outer bound for the quadratic Gaussian case was found by Ozarow [2]. Practical ways to create multiple description from a single description has been suggested by many. Some of the ways are Unequal Error Protection (UEP) [3] ,Correlating Transforms [4], Dithered Delta-Sigma Quantization [5] and Multiple Description Scalar quantization (MDSQ) [6].

While MD coding handles the situation where the rate in some channels drops to zero temporarily, thus causing packet-loss, it fails to accommodate more subtle changes in link rate such as rate reduction. In such a case, a classic MD scheme can not use the link capacity left for information transfer, causing even minor rate reduction to be considered as link failure. In order to accommodate such a frequent situation, we have to add the quality of Successive Refinability (SR) to the description. A source is successively refinable if encoding in multiple stages incurs no rate loss as compared with optimal rate-distortion encoding at the separate distortion levels. Therefore, an SR description can be downgraded to a lower rate by extraction of previous stages from the original description. Furthermore, we have to add a ”transcoding” functionality to the nodes that the description travels through. They can not only either pass or drop the description, depending on sufficient rate, but have the ability to manipulate it (e.g. by extracting some previous stages of SR description). By this we extend the Multiuser Successive Refinement (MSR) problem, formulated by Pradhan and Ramchandran in [7] and explored in [8], so it would treat uncontrolled packet-loss, as well as controlled successive refinement.

Summing up, we propose a more modular scheme for transmitting over diversity system, which have the following qualities:

1. The description is successively refinable - unexpected reduction in link's rate can be solved by downgrading the original description into a more coarse description, so it would fit to the new link's rate.
2. Work in a distributed environment - each node of the network can downgrade the description, relaying only on local information, i.e. knowing neither the other descriptions, nor the entire network topology.
3. Maintain the aforementioned original MD properties

We call such scheme a Multi Level Multiple Description (MLMD) scheme. Based on Vaishampayan's seminal work on MDSQ [6], we present two constructive methods of building such an MLMD scheme: *Deflated MDSQ* which prioritize the maximal rate, and *Inflated MDSQ* which prioritize the minimal rate.

The paper is organized as follows: Relevant rate distortion bounds from the literature are presented in Section 2. In Section 3 MDSQ is presented and analyzed. In Section 4 we present the design of MLMD, and analyze it in Section 5. Performance results are presented in Section 6. Section 7 discuss open questions and concludes.

## 2. RATE DISTORTION BOUNDS

El Gamal and Cover [1] found the achievable rate-distortion region for the MD problem, and Ozarow [2] showed that for the Gaussian source and squared-error distortion measure, this region is indeed tight.

For the situation depicted in fig.1(a), Ozarow showed that for a unit variance source given $R_1$ and $R_2$ any triple of average distortions $\{D_0, D_1, D_2\}$ (associated with the respective reconstructions) could be achieved if and only if:

$$\begin{aligned} D_1 \geq 2^{-2R_1}, D_2 \geq 2^{-2R_2} \\ D_0 \geq \frac{2^{-2(R_1+R_2)}}{1-(\sqrt{\Pi}-\sqrt{\Delta})^2} \end{aligned} \tag{1}$$

where $\Pi = (1-D_1)(1-D_2)$ and $\Delta = D_1 D_2 - 2^{-2(R_1+R_2)}$. Consequently, Vaishampayan [9] proved that for the symmetric case ($R_1 = R_2 = R$)

$$\begin{aligned} D_1 = D_2 = D_{1,2} = \Theta(2^{-2R(1-\alpha)}) \\ D_0 = \Theta(2^{-2R(1+\alpha)}) \end{aligned} \tag{2}$$

for some $1 \geq \alpha \geq 0$, and found an equivalent criteria for optimality, by inspecting the distortion product

$$D_0 D_{1,2} = \Theta(2^{-4R}). \tag{3}$$

The optimality is in the asymptotic sense, thus neglecting constant factors such as scalar vs. vector quantization loss.

## 3. MULTIPLE DESCRIPTION SCALAR QUANTIZATION

Vaishampayan [6] suggested an MD scheme based upon scalar quantization (MDSQ). Each description is a union of disjoint intervals. The problem of cleverly mapping those intervals (represented by the indices of the scalar quantizer) to descriptions, is called the index assignment problem. An easy way to visualize a two-dimensional MDSQ, is by using an index assignment matrix as depicted in fig.2(a). One description consists of the vertical coordinate selecting the row, and the other consists of the horizontal coordinate selecting the column. The encoder quantize the source, using a scalar quantizer, to an index value, then sends its row and column coordinates along different paths as descriptions. If the decoder gets both descriptions, it can find the original cell by intersecting the row and the column, and then reconstruct the value associated with this index.If only one description makes it to the decoder, it knows only the row or column of the original cell. In that case, its reconstruction is the centroid of all the values associated to indices contained in the row or column.

Vaishampayan describes a method to build such matrix which he calls nested index assignment. The matrix in Fig.2(a) was created using that method. Following Vaishampayan we'd use the following definitions: n is the number of rows or columns, k is the number of symmetric matrix diagonals that are filled with indices, codewords (or central cells) are the total number of indices and spread is the difference between the lowest index to the highest index in a certain group of indices. $spread = max(index) - min(index) + 1$

It's clear that the spread of a set of indices determines the quality of the reconstruction produced from them. The bigger the spread, the bigger the distortion is. Vaishampayan proved that using this method, the number of codewords is

$$codewords = (1 + 2k)n - k(k + 1) \tag{4}$$

and the maximal spread is

$$spread = 2k^2 + k + 1 \tag{5}$$

We can see that $k$ is the tradeoff controlling parameter. By increasing $k$ the number of codewords raises, meaning a better central distortion($D_0$), but it simultaneously increases the spread, meaning worse side distortion($D_{1,2}$).

Vaishampayan proved that if the size of every quantizer's cell is

$$\Theta(1/N^b); 1 \leq b \leq 1 \tag{6}$$

then the corresponding MSE is

$$\Theta(1/N^{2b}) \tag{7}$$

Substituting equations 4,5, into 6,7 yields that the central distortion and side distortion are

$$D_0 = \Theta([(1 + 2k)n - k(k + 1)]^{-2}) = \Theta((nk)^{-2})$$

$$D_{1,2} = \Theta([\frac{2k^2 + k + 1}{(1 + 2k)n - k(k + 1)}]^2) = \Theta((n/k)^{-2}) \tag{8}$$

Remembering that $n = 2^R$ and setting $k = 2^{R/n}$

$$D_0 = \Theta(2^{-2R(1+1/n)})$$

$$D_{1,2} = \Theta(2^{-2R(1-1/n)}) \tag{9}$$

thus making the distortion product to be

$$D_0 D_{1,2} = \Theta(2^{-4R}) \tag{10}$$

proving that this is indeed an optimal distortion pair, in the sense of (3).

## 4. DESIGNING A MULTI LEVEL INDEX ASSIGNMENT

As mentioned previously, in an MLMD system, suited for real-world packet switched network, a new type of node is introduced - the transcoder. The transcoder has the ability to manipulate the payload of the packet, besides possible other roles such as switching or decoding. In the situation depicted in fig.1(b) , if $R_1 > R_4$ then transcoder1 has to decrease the rate of the forward transmitted description. The transcoder has to do it relying only on local information. In order to do so, all the other parts of the system (encoder, decoders) have to be designed to enable the transcoding.

### 4.1. Encoder design

The encoder acts very much like the original MDSQ encoder, the changes are denoted in italics font. Its descriptions are the coordinates of an index in the *MLMD matrix*. It aggregates several instances of a description over time , and sends them as a packet. It sets the packet header to contain the number of description instances sent, their *resolution* (that may change during the packet's journey, due to its MLMD nature) and dimension (in a two dimensional MD system this parameter may be either vertical or horizontal). Since a packet is composed from many instances of the description, the size of the header is negligible.

Besides having the resolution parameter in the header, the main change in the MLMD encoder is the construction of its index assignment matrix. We suggest two possible approaches to construct it:

*Inflated MLMD* - Create an initial index assignment matrix using original scheme. Set its rate (by setting $n$ to an appropriate value) to be the minimal rate we want to deal with. Now, inflate each matrix cell to the maximal rate. Say the original matrix is A, the inflated is B, and the Inflating Factor (IF) is defined as follows:

$$IF = 2^{MaximalRate - MinimalRate}, \tag{11}$$

then

$$B[i][j] = \begin{cases} (A[i/IF, j/IF] - 1)IF^2 + 1 & \text{(i mod IF)=(j mod IF)=0} \\ \dots & \\ (A[\lfloor i/IF \rfloor, \lfloor j/IF \rfloor])IF^2 & \text{(i mod IF)=(j mod IF) =IF - 1} \end{cases}$$

In section 5 we will prove, that only the smallest and largest index of a row or column matter, hence the inner mapping of the other inflated indices is not described explicitly. For a possible mapping, See Fig.2(b).

Since inflated MLMD is based upon a lower (minimal) rate matrix, it gives priority to the lower rate, in the sense that the designer can control the tradeoff of side to central distortion for the lower rate directly, and the tradeoffs of the higher rates are derived from it.

*Deflated MLMD* - Create an initial index assignment matrix using original scheme. Set its rate to be the maximal rate. In section 5 we will prove that the original MDSQ matrix is also an MLMD matrix. Since deflated MLMD is based upon a higher (maximal) rate matrix, it gives priority to the higher rate, in the same sense that was described earlier for the inflated MLMD.

### 4.2. Transcoder design

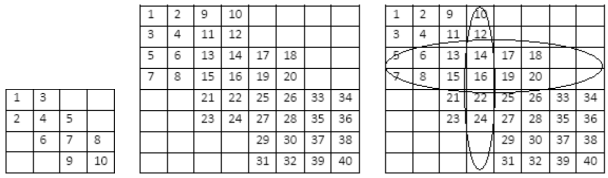
The transcoder transmits the description. In the case it does not have enough rate to send the description tuple, it chops the Least Significant Bits (LSB) from the description, until the rate limit is satisfied. It updates the resolution field in the header appropriately. The transcoder is very simple, and totally unaware to the MLMD nature of the transmission, it just chops bits.

### 4.3. Decoder design

As in the case of the encoder, the MLMD decoder is very similar to the MDSQ one. Each description received selects a set of indices from the matrix. the decoder intersects these sets to construct the final set. The reconstruction it announces is the centroid of all the values associated to the indices in the final set. The main difference from the MDSQ decoder, is that an MLMD description can choose a set that consists of *couple* of the matrix rows (or columns) and not just one of them. The decoder first calculates the factor of the description $factor = \frac{n}{2^{resolution}}$ ,where n is the MLMD matrix size, and resolution is the *actual* rate of the description, known to the decoder from packet header. The set induced by the description is all the indices starting from the $(factor * description)$ row / column, to the $(factor * (description + 1) - 1)$ row / column. Note that the decoder is totally oblivious to the method in which the matrix was created - same decoder for inflated or deflated MLMD matrix.

### 4.4. Illustration

The whole picture may be clarified using an example. In the situation depicted in fig.1(b), say $R_1 = R_2 = R_5 = 3$ bits and $R_4 = 2$ bits. The encoder wants to send the index 16. Using the high resolution matrix it sends its coordinates - 3,3 (zero based addressing, encoded as binary strings of 3 bits length - ”011”) as description 1 and 2 to decoder 1 and 2, respectively. Since $R_2 = R_5$ decoder 2 can send its description directly to decoder 5 and 3, but transcoder 1 have to do some transcoding operation so it can send the 3 bit wide description 1 on a 2 bit rate channel. Therefore, it chops the last bit of the description and transform it from ”011” to ”01”. The decoding of decoder 3 is shown in fig.2(c). When it receives both descriptions it first finds the relevant index set for both descriptions: the third and fourth row for the chopped description 1 and the fourth column for description 2. After both sets are intersected with eachother only $\{14,16\}$ remain. The reconstruction is the centroid of the values associated by the quantizer to indices 14 and 16.




(a) (b) (c)

**Fig. 2**. Index Assignment Matrices: (a) original matrix n=4, k=1, codewords=10, max spread=4; (b) inflated matrix based on (a), IF=2; (c) MLMD decoding example

## 5. ANALYSIS

We will analyze the behavior of the distortion of our algorithm in the same manner and under the same assumptions Vaishampayan analyzed his (see Section 3). Although the MLMD system can handle asymmetric rate reduction, as shown in Sec.4.4, for brevity only the symmetric rate reduction is analyzed.

For the inflated version, we note that if the maximal spread for the original low resolution matrix was $spread_{lr} = MaxIndex_{lr} - MinIndex_{lr} + 1$, then by using the inflate function from Section 4.1 the corresponding spread is bounded by

$$\begin{aligned} spread &= MaxIndex_{hr} - MinIndex_{hr} + 1 \\ &\leq IF^2 MaxIndex_{lr} - [IF^2(MinIndex_{lr} - 1) + 1] + 1 \\ &= IF^2(MaxIndex_{lr} - MinIndex_{lr} + 1) = IF^2(spread_{lr}) \end{aligned} \tag{12}$$

and the number of the codewords is

$$codewords = IF^2(codewords_{lr}). \tag{13}$$

Therefore, the central and side distortions are (using the assignments from (9))

$$\begin{aligned} D_0 &= \Theta[\frac{1}{codeswords}]^2 = \Theta([IF^2 codewords_{lr}]^{-2}) = \\ &\Theta(IF^{-4} D_{0;lr}) \\ D_{1,2} &= \Theta([\frac{spread}{codeswords}]^2) = \Theta([\frac{IF^2 spread_{lr}}{IF^2 codeswords_{lr}}]^2) = \\ &\Theta([\frac{spread_{lr}}{codeswords_{lr}}]^2) = D_{1,2;lr} \end{aligned} \tag{14}$$

which yields distortion product of

$$\begin{aligned} D_0 D_{1,2} &= D_{0;lr} D_{1,2;lr} IF^{-4} = \\ &\Theta(2^{-4R_{lr}} 2^{-4(R-R_{lr})}) = \Theta(2^{-4R}) \end{aligned} \tag{15}$$

thus proving the description optimality, in the sense of (3). We can see that using this method, the central distortion decays rapidly as a function of the rate, while the side distortion remains constant [1].

[1] As a by-product, we have found another optimal index assignment method for the MDSQ system, which

**Fig. 3**. Applying MLMD on a picture: $R_1 = R_2 = 4$, $R_3 = R_4 = 2$. Deflated matrix $n = 16, k = 2$. Reconstructions (from left to right) : $x_0, x_1, x_2, x_3, x_4, x_5$

For the deflated version, first we'd take a look on the construction of the original index assignment matrix. It's constructed from r-shaped parts, their center on the main diagonal, each part holds $2k + 1$ successive indices. Therefore adding a portion of a new r shaped part to an index set, adds at most $2k + 1$ to its spread. By chopping bits, more rows (for a vertical coordinate description) to the description set, each row adds one new r shaped part. The number of rows is $IF$, so there are $IF - 1$ new rows. If we add now $IF - 1$ new columns to the set it would not change the set spread, since no new r shaped parts are added. That means we have to add $\hat{spread} = (IF - 1)(2k + 1)$ to the numerator of both side and central distortion. Therefore, the central distortion (the distortion associated with reconstruction $x_3$, as depicted in fig1(b)) is:

$$D_3 = \Theta([\frac{1 + \hat{spread}}{codeswords}]^2) = \Theta([\frac{(IF-1)(2k+1)+1}{(1+2k)n - k(k+1)}]^2) = \Theta([\frac{IF}{n}]^2) = \Theta([\frac{2^{R_{hr}-R_{lr}}}{2^{R_{hr}}}]^2) = \Theta(2^{-2R_{lr}}) \tag{16}$$

and the side distortion is

$$\Theta([\frac{spread + \hat{spread}}{codeswords}]^2) = \Theta([\frac{2k^2 + k + 1 + (IF-1)(2k+1))}{(1+2k)n - k(k+1)}]^2).$$

If $IF >> k$, then $2k^2 + k + 1$ is negligible in comparison to $(IF - 1)(2k + 1) + 1)$ and the equation becomes

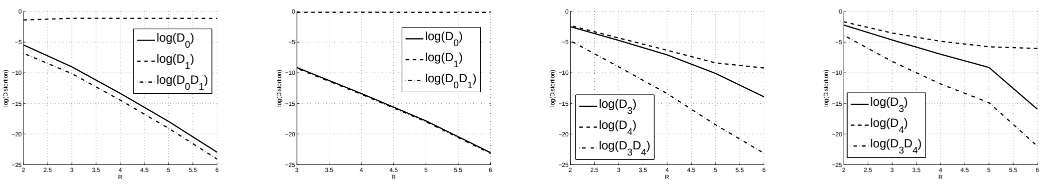
$$= \Theta([\frac{(IF-1)(2k+1)+1}{(1+2k)n - k(k+1)}]^2) = D_0 = \Theta(2^{-2R_{lr}}) \tag{17}$$

which yields distortion product of

$$D_3 D_{4,5} = (D_3)^2 = \Theta(2^{-4R_{lr}}). \tag{18}$$

This proves the description is optimal, in the sense of (3), provided $IF >> k$ - which happens either when the central to side distortion ratio is close to 1 (k is small), or when the rate reduction (denoted by IF) is large. We can see that using this method the side and central distortion in (18) decay at the same rate.

offers more working points for the designer. While in the original index assignment the number of diagonals ($k$) may take only an integer value, using the inflated method we can create MDSQ with punctured diagonals, i.e. with a fractional k.




(a) (b) (c) (d)

**Fig. 4**. Performance results: (a) Inflate, original matrix n=4, k=1. (b) Inflate, original matrix n=8, k=3. (c) Deflate n=64, k=1. (d) Deflate n=64, k=3.

## 6. PERFORMANCE

We've tested our algorithm by simulating the system in fig.1(b). The source is Gaussian noise $f(x) \sim N(0, 1)$. We trained our scalar quantizer to it, using the max-Lloyd algorithm. Then, we encoded the index received from the quantizer as descriptions, using the MLMD matrix (created by the inflate/deflate method) , passed the descriptions through transcoders, and used the decoder described in section 4.3 to reconstruct the original value. Finally, we'd compared the reconstruction to the original signal, using squared-error distortion measure, and plotted a graph of $\lg_2 distortion$ (so it would be easy to see that the slope is indeed -4) as a function of the rate (in bits). The results appear in fig.4.As expected, the inflated method simulations shown at figs.4(a),4(b) achieves a slope of $-4R$ for the distortion product. Only the central distortion ($D_0$) contributes to the slope, while side distortion ($D_1$) maintain its values regardless to the rate.

The deflated method simulations shown at figs.4(c),4(d) achieves a slope of $-4R$ for the distortion product, and a slope of $-2R$ for both central and side distortion, only when $k$ is small as in fig.4(c) or when the rate reduction is large - the lower rates of fig4(c) and fig4(d).

For illustrative purposes, we applied the same process to a picture, and the results are shown in fig.3.

## 7. CONCLUSIONS

The problem of rate reduction in an MD system is introduced. A Multi Level MD (MLMD) approach is presented, and a design algorithm based upon the MDSQ method for the encoder, decoders and transcoders of the system is suggested. The algorithm is analyzed and proved to be asymptotically optimal, and the simulation results obtained show its optimality.

While in this paper we considered MLMD based upon MDSQ, it's still an open question whether MLMD can be based also upon other methods of MD. Since our method rely heavily on the exact details of the coding of the specific MD system, it can not be deployed "out of the box" for another MD system. However, it seems that the underlying principles can be applied with appropriate changes to any specific MD method. For instance, consider the MDCT (MD Correlating Transform) system [4]. In that system the description is created by sending a linear transform of two instances of the source. If both descriptions make it to the decoder, it can solve the linear equations and reconstruct the original information. In the case that only one of the descriptions arrives, the decoder can estimate the other description and then extract the information. It's clear that if the description is changed a little by chopping the LSB, it would cause only a little change in the reconstructed data. Thus, MLMD system can be based upon MDCT system. Even for more complex MD systems such as dithered delta-sigma quantization MD [5] it seems that if we'd add successive refinability to the description by using SR quantizer inside the noise shaping loop, then it could be used for MLMD encoding.


# Acknowledgement

The authors would like to thank J. Øestergaard for providing useful comments on a draft version of this paper.


# References


[1] A.A. El Gamal and T.M. Cover, "Achievable rates for multiple descriptions", *IEEE Trans. Inform. Theory, vol. 28*, pp. 851-857, Nov. 1982

[2] L. Ozarow, "On a source-coding problem with two channels and three receivers", *Bell Syst. Tech. J., vol. 59, no. 10*, pp. 1909-1921, Dec. 1980.

[3] R. Puri and K. Ramchandran, Multiple description source coding using forward error correction, *in Conf. Rec. 33rd Asilomar Conf. Sig., Sys., & Computers, Pacific Grove, CA, Oct. 1999, vol. 1*, pp. 342-346.

[4] V.K. Goyal and J. Kovacevic, "Generalized multiple description coding with correlating transforms," *IEEE Trans. Inform. Theory, vol. 47*, Sept. 2001.

[5] Jan Ostergaard and Ram Zamir, "Multiple-Description Coding by Dithered Delta-Sigma Quantization," *Proceedings of the 2007 Data Compression Conference*, pp 63-72,March 2007

[6] V.A. Vaishampayan, "Design of multiple description scalar quantizers", *IEEE Trans. Inform. Theory, vol. 39*, pp. 821-834, May 1993.

[7] S. Pradhan and K. Ramchandran, "Multiuser successive refinement," *in Conference on Information Sciences and Systems*, Mar. 2002.

[8] C. Tian , J.Chen, S.N. Diggavi "Multiuser Successive Refinement and Multiple Description Coding"," *IEEE Trans. Inform. Theory, vol. 54*, pp. 921-931, Feb. 2008.

[9] V.A. Vaishampayan and J.-C. Batllo, "Asymptotic analysis of multiple description quantizers", *IEEE Trans. Inform. Theory, vol. 44*, pp. 278-284,Jan. 1998.