

MULTIPLE DESCRIPTION CODING VIA SCALING-ROTATION TRANSFORM

Wenqing Jiang and Antonio Ortega

Integrated Media Systems Center
 Department of Electrical Engineering-Systems
 University of Southern California
 Los Angeles, CA 90089-2564
 E-mail: {wjjiang, ortega}@sipi.usc.edu

ABSTRACT

In this paper, we propose a two-stage transform design technique for Multiple Description Transform Coding. The first stage is the structure design in which we enforce a Scaling-Rotation factorization of the transform and we further constrain the transform for specific channel conditions using the knowledge of the input correlation matrix and the desired output correlation matrix. In the second stage, magnitude design, we find the optimal transform from all admissible transforms given by the structure design using the numerical algorithm proposed by Goyal et al. [1]. Such a design enables a structured transform framework which reduces both the design and implementation complexities compared to an exhaustive search through the whole space of nonorthogonal transforms. We give two examples to illustrate the design idea, the Scaling-Hadamard transform for equal rate channels and the Scaling-DST transform for sequential protection channels.

1. INTRODUCTION

Multiple Description Coding (MDC) has recently been shown to be effective as a diversity technique for robust communication over erasure channels [2, 3, 4]. The problem of two-description coding was first formulated by El Gamal and Cover [5]. Different MDC systems have since been proposed in the literature, which include the design using a Multiple Description Scalar Quantizer, (MDSQ)[6, 3] and the design using a correlating transform [2, 7, 1, 4].

In this paper we study a Multiple Description Transform Coding (MDTC) system as shown in Fig. 1 [8]. The input data is first decorrelated using T_1 (e.g. KLT/DCT). After quantization, another transform T_2 is applied to recorelate the data which is split over different channels for transmission. Since there exists correlation between descriptions, the receiver can estimate the lost descriptions from received descriptions in case of channel failures. The design of a MDTC system focuses on the search for optimal redundancy rate-distortion points by designing the correlating transform T_2 . In the sequel, we consider only the design of T_2 .

For a pair of Gaussian random variables with two output channels, Orchard et al. [7] and Goyal et al. [1] have shown that nonorthogonal transforms perform better compared to orthogonal transforms in terms of redundancy rate-

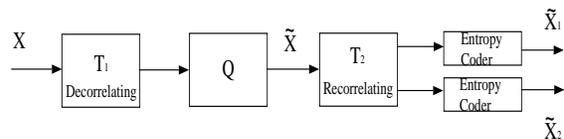


Figure 1: A typical MDTC system

distortion gain. For MDTC systems with M inputs and M outputs, the transform design and its performance analysis is still an open problem. Orchard et. al [7] suggested a redundancy allocation strategy among pairs of input variables but optimal pairing is not yet readily available. A numerical optimization algorithm was proposed by Goyal et al. [1] to design transforms for arbitrary number of channels. However, exhaustive search through the whole space of all nonorthogonal transforms is not only computationally intensive but also leads to implementation difficulties when using an arbitrarily structured nonorthogonal transform.

In this paper, we propose a two-stage transform design technique for Multiple Description Transform Coding, i.e., structure design and magnitude design. The motivation is that protection properties of a MDTC system can be characterized by the output correlation matrix, i.e., which descriptions are correlated (structure) and to what extent they are correlated (magnitude). While the magnitude information can not, in general, be quantified for specific redundancy and distortion constraints, the structural information can be inferred from specific channel conditions or protection requirements. Consequently, the structure design will find admissible transforms (eigenmatrices of the output correlation matrix) using a Scaling-Rotation (SR) factorization and the magnitude design will search for the optimal transform from these admissible transforms. Such a design enables a structured transform framework and reduces both the design and implementation complexities because we have imposed some structures on the space of transforms to be searched.

The remainder of this paper is organized as follows. In Section 2, we give the two-stage transform design approach based on parametric Scaling-Rotation transforms for the design of MDTC systems with M inputs and M outputs. In Section 3, we give two design examples for equal rate channels and sequential protection channels. Simulation results of Gaussian vectors and analysis of MDTC systems are given in Section 4. We conclude our work in Section 5.

2. PROPOSED TRANSFORM DESIGN APPROACH FOR MDTC

For MDTC of pairs of independent Gaussian random variables, an important result is that, nonorthogonal transforms are better than orthogonal transforms in terms of the redundancy rate-distortion gain [7, 1]. For a given central distortion, the nonorthogonal transform not only achieves lower average side distortion using same amount of extra bits, it can also extend the redundancy rate-distortion function to the region that the orthogonal transform can not reach. However, nonorthogonal transforms pose two challenges to the MDTC system design: (1) quantization has to be performed before the transform which requires that the transform has to be an integer transform (mapping integers to integers); (2) the problem of optimal transform design based on the numerical algorithm by Goyal et al. [1] becomes computationally intensive with the increase of M since the number of parameters increases quadratically ($O(M^2)$ for a $M \times M$ transform); and (3) the numerical algorithm [1] may lead to an optimal transform arbitrarily structured which increases implementation difficulties. The next two sections will study these issues in turn.

2.1. Parametric Transforms and Factorizations

To solve the first problem, we will design the transform directly from its QR factorization [9]. This factorization exists for any matrix with linearly independent columns, either square or nonsquare. We mention that nonsquare transforms are special cases of frame expansions which have also been used to add redundancy for MDTC [4]. To reduce the design complexity, we only consider square transforms in this work.

Using our notations we write QR factorization as $T = RU$ where R is an orthogonal transform and U is an upper triangular matrix. In fact, the upper triangular matrix U can be further decomposed as a product of a scaling matrix S and another lifting matrix L (an upper triangular matrix with diagonal entries all 1). In other words, any transform can be written as a concatenation of three transforms, a Lifting transform L , a Scaling transform S and a Rotation transform R (orthogonal transform), i.e., $T = RSL$. To reduce the design complexity, we only study SR transforms in this work, i.e., $T = SR$. To guarantee lossless factorization, we have to impose the constraint that the determinant of T has to be set to one, i.e., $\det(T) = \det(SR) = \det(S) = 1$.

Next we show SR transforms can be factored into lifting steps hence can be implemented losslessly. We first present factorization results for 2x2 rotation and scaling transforms (adapted from [10]).

$$\begin{aligned} \begin{bmatrix} a & 0 \\ 0 & 1/a \end{bmatrix} &= \begin{bmatrix} 1 & a - a^2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1/a & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & a - 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \\ \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} &= \begin{bmatrix} 1 & (\cos \theta - 1)/\sin \theta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \sin \theta & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & (\cos \theta - 1)/\sin \theta \\ 0 & 1 \end{bmatrix} \end{aligned}$$

It is well known [9] that any orthogonal $M \times M$ matrix Q_M can be factored into the product of $M(M-1)/2$ orthogonal matrices, each of which is a $M \times M$ Givens rotation which only rotates two components at a time. This indicates that any orthogonal transform can be factorized into lifting steps using the result of 2x2 rotation transform. Factorization of a scaling transform is also very straightforward since we require its determinant to be 1. It can be shown that such $M \times M$ scaling transform can be written as a concatenation of $M-1$ subsampling transforms, each of which only scale two input variables at a time. Hence we can use the factorization result of a 2x2 scaling matrix. Thus we have shown that any SR transforms can be factored into lifting steps hence can be implemented losslessly. The parametric Scaling-Rotation (SR) transform assumes the form

$$T(\mathbf{A}, \Theta) = R(\Theta)S(\mathbf{A}) \quad (1)$$

where \mathbf{A} are scaling factors and Θ are Givens rotation angles.

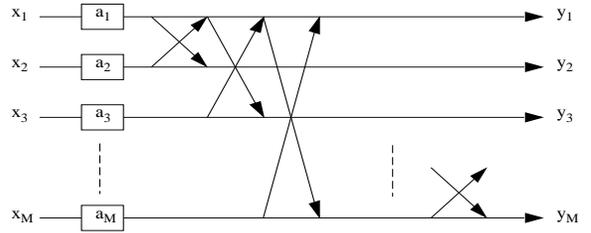


Figure 2: Lattice structure of a SR transform

The general structure of a SR transform is shown in Fig.2 which has a lattice structure similar to that of a biorthogonal filterbank. The perfect reconstruction property is guaranteed by such a transform framework. To derive the inverse transform, one only need to change all the rotation angles to the opposite sign and multiply by the inverses of the scaling factors. Using such a SR transform framework, we can also design the integer transform directly rather than obtaining the integer transform from lifting factorization of the designed non-integer transform. Thus we can completely eliminates the lifting factorization stage and helps to improve the system design accuracy by avoiding unnecessary rounding errors.

2.2. Two Stage Transform Design

Although SR transforms enjoy structured lattice implementations, the number of design parameters of an $M \times M$ SR transform still increases quadratically ($O(M^2)$). In this section we propose a two-stage design technique making use of the available channel information to further constrain the SR transform and reduce both the design and implementation difficulties.

Let $R_X = \text{diag}\{\sigma_{ii}^2\}$, $i = 1, 2, \dots, M$ be the correlation matrix of input vector X . Let $R_Y = \{r_{ij}\}$, $i, j = 1, 2, \dots, M$ be the correlation matrix of the transform output Y . Then we have

$$R_Y = T(\mathbf{A}, \Theta)R_X T(\mathbf{A}, \Theta)^t \quad (2)$$

$$= R(\Theta)R_S R(\Theta)^t \quad (3)$$

where $R_S = S(\mathbf{A})R_X S(\mathbf{A})^t = \{a_{ii}^2 \sigma_{ii}^2\}, i = 1, 2, \dots, M$ is a diagonal matrix. We can see that the rotation transform $R(\Theta)$ has to be the eigenmatrix of the required output correlation matrix R_Y . If the output correlation matrix can be pre-designed, then the problem of correlating transform design can be formulated as the inverse problem of matrix diagonalization. An orthogonal solution to the correlating transform is simply the inverse of KL transform. In general, we can not pre-design the output correlation matrix subject to redundancy and distortion constraints, however, we can infer its structural information for specific channel conditions or protection requirements. For example, equal rate channels require the output correlation matrix to have equal diagonal entries for Gaussian inputs. Thus admissible transforms have to be able to generate correlation matrices with equal diagonal entries.

Based on such observations, we propose a two stage transform design approach, i.e., structure design and magnitude design. The structure design finds admissible transforms (eigenmatrices of the output correlation matrix) for specific channels using the Scaling-Rotation (SR) factorization framework. In Fig.3 we show that the transform search space can be reduced gradually using available channel information (details in next section). The magnitude design then searches for the optimal transform from admissible transforms using the algorithm described in [1] where derivation details of the average side distortion $D_s(\mathbf{A}, \Theta)$ and the redundancy bit rate $\rho(\mathbf{A}, \Theta)$ are given. We perform a redundancy constrained transform design using a Lagrangian multiplier λ . The cost function is $J = D_s(\mathbf{A}, \Theta) + \lambda \rho(\mathbf{A}, \Theta)$. By varying λ , one can scan all the operational redundancy rate distortion points (D_s, ρ) .

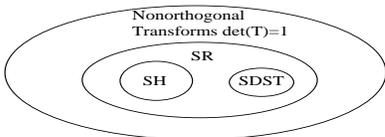


Figure 3: Transform search space.

3. MDTC DESIGN EXAMPLES

We give design examples for two important channels, equal rate channels and sequential protection channels, both of which can be characterized by the output correlation matrix R_Y . It turns out for these two special channels, not only can we use fixed rotation transforms but also these fixed transforms have fast algorithms. Using a fixed rotation transform, we can reduce the number of design parameters from $M - 1(\mathbf{A}) + M(M - 1)/2(\Theta) = (M^2 + M - 2)/2$ to only $M - 1$. This makes the optimization converge faster and reduces the amount of information to be conveyed to the decoder. On the other hand, fast algorithms reduce both the encoding and decoding complexities.

3.1. Equal Rate Channels

The equal rate channels requires that output descriptions have same rates which helps the buffer management (e.g. packetization/depacketization in a packet network) both at

the encoder and the decoder. We mention that equal rate should be interpreted in a statistical sense. For example, two Gaussian sources with same variances will be viewed as equal rate sources if quantized with the same quantizer. That is to say, for Gaussian random variables, “equal rate” requires that the R_Y to have equal diagonal entries. Here we provide a SR transform which generates equal rate descriptions for arbitrary number of channels with $M = 2^n$.

The equal rate transform we propose is a Scaling Hadamard transform.

$$T(\mathbf{A}) = HS(\mathbf{A}) \quad (4)$$

We need to show, under this transform, for input X , R_Y has equal diagonal entries (equal variances), i.e., $r_{ii} = r, i = 1, 2, \dots, M$ and r is a constant.

Since the Hadamard transform H is real symmetric, the output correlation is $R_Y = HR_S H$. Denote $H = \{h_{ij}\}, i, j = 1, 2, \dots, M$. We have $h_{ji}^2 = 1, \forall (i, j)$. The output Y component variance r_{ii} can be written as

$$r_{ii} = \sum_{j=1}^M a_{jj}^2 \sigma_{jj}^2 h_{ji}^2 = \sum_{j=1}^M a_{jj}^2 \sigma_{jj}^2$$

Thus we have shown a Scaling-Hadamard (SH) transform generates equal rate descriptions. We mention that this Scaling-Hadamard transform reduces to the optimal equal rate transform given by Goyal and Kovacevic [1] when $M = 2$ which demonstrates that, at least for two descriptions coding, SR transform does not compromise the optimality of the MDTC system. We also note that the cascaded structure given by Goyal et al. [1] is equivalent to a Scaling Hadamard transform for $M = 4$.

3.2. Sequential Protection Channels

Another example channel is the sequential protection channel in which descriptions are sent out sequentially and each description will only protect the losses of its immediate predecessor and its immediate successor. In a lossy packet network, an example scenario is that when each packet carries information for the recovery of its previous and next packets, e.g. a similar case when Robust Audio Tool technique is applied for audio transmission [11]. For a MDTC system, this indicates that the output correlation matrix R_Y should be a tridiagonal matrix in which descriptions are sequentially correlated. The R_Y assumes the form

$$R_Y = \begin{bmatrix} 1 & \rho & 0 & 0 & \dots & 0 \\ \rho & 1 & \rho & 0 & \dots & 0 \\ 0 & \rho & 1 & \rho & \dots & 0 \\ & \vdots & & \vdots & & \\ 0 & \dots & 0 & \rho & 1 & \rho \\ 0 & \dots & 0 & 0 & \rho & 1 \end{bmatrix} \quad (5)$$

From matrix theory, we know that the eigenmatrix for this type of symmetric tridiagonal Toeplitz matrices is the Discrete Sine Transform (DST) [12]. So the transform we propose for sequential protection channels is the Scaling-DST transform

$$T(\mathbf{A}) = DST S(\mathbf{A}) \quad (6)$$

4. SIMULATION RESULTS AND DISCUSSIONS

In this experiment, we compare results of different configurations of the correlating transform for a 4D Gaussian vector source with standard deviations $\{1, 0.5, 0.3, 0.1\}$ [7]. We compare the side distortion when there is only one description lost with equal channel failure probabilities. The different transforms are (i) Rotation; (ii) Scaling-Rotation; (iii) Scaling-Hadamard; and (iv) Scaling-DST. The configuration Scaling-Hadamard is equivalent to the cascaded structure given by Goyal and Kovacevic (Fig.3 in [1]) for a 4-D input vector. The optimization is done via Powell's direction set technique [13]. The initial scaling factors are all set to be 1s and initial rotation angles are all set to be $\pi/4$ for the Scaling-Rotation configuration.

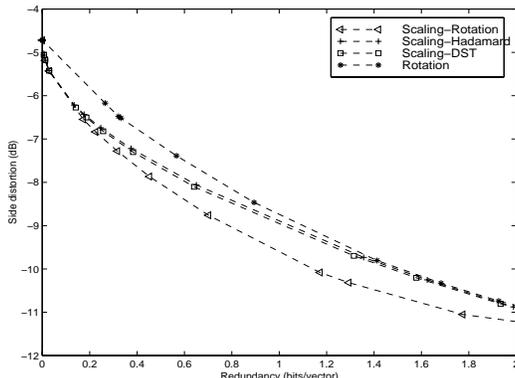


Figure 4: Comparisons among different transforms.

The comparison of all four configurations is shown in Fig.4. Clearly all the nonorthogonal transforms achieve better performances compared to the orthogonal transform (case (i)). This indicates that nonorthogonal transforms can also perform better than orthogonal transforms for MDTC systems of more than two channels. We also observe performance degradations when we impose constraints on the rotation transform, Scaling-Hadamard or Scaling-DST transform, specially at higher redundancy bit rates. However, using structured transforms, we reduce the transform parameters necessary to be delivered to the decoder and we also reduce the implementation complexities. It is interesting to note that the Scaling-Hadamard transform has approximately the same performance as the Scaling-DST transform when there is a single description lost.

However, the disadvantage for a complete MDTC system design (refer Fig.1) is that, using a nonorthogonal transform necessitates a framework which needs two transforms, T_1 for decorrelation (before quantization) and T_2 for recorelation (after quantization), both at the encoder and decoder. Obviously, this increases the system implementation complexity. If an orthogonal transform is used, then one can merge T_2 with T_1 into one transform to reduce the computation. Such a single transform can decorrelate and recorelate simultaneously the input vector and has been studied in the context of vector quantization as Karhunen-Loeve Vector Transform (KLVT) [14]. Such an alternative to the MDTC system design provides a tradeoff between redundancy-rate distortion gain and the computational complexity which may be useful for delay-constrained coding applications such as streaming-video over Internet.

5. CONCLUSIONS AND ACKNOWLEDGMENTS

In this paper, we have proposed a two-stage transform design approach, structure design and magnitude design, for MDTC systems using SR transforms. Such an approach enables us to find structured transform solutions using available channel information thus reduces both the design and implementation complexities. Examples are given for equal rate channels using SH transforms and for sequential protection channels using SDST transforms.

The authors would like to thank Jelena Kovacevic, Vinay Vaishampayan and Sergio Servetto for helpful discussions and for providing the preprints [4, 3].

6. REFERENCES

- [1] V. K. Goyal and J. Kovacevic, "Optimal multiple description transform coding of gaussian vectors," in Proc. of IEEE Data Compression Conference, 1998.
- [2] Y. Wang, M. Orchard, and A. R. Reibman, "Multiple description image coding for noisy channels by paring transform coefficients," in Proc. IEEE 1997 First Workshop on Multimedia Signal Processing, 1997.
- [3] S. D. Servetto, K. Ramchandran, V. Vaishampayan, and K. Nahrstedt, "Multiple description wavelet based image coding," in ICIP'98, 1998.
- [4] V. K. Goyal, J. Kovacevic, R. Arean, and M. Vetterli, "Multiple description transform coding of images," in ICIP'98, 1998.
- [5] A. A. El-Gamal and T. M. Cover, "Achievable rates for multiple descriptions," IEEE Trans. Information Theory, vol. IT-28, no. 6, pp. 851-857, Nov. 1982.
- [6] V. A. Vaishampayan, "Design of multiple description scalar quantizers," IEEE Trans. Information Theory, vol. 39, no. 3, pp. 821-834, 1993.
- [7] M. T. Orchard, Y. Wang, V. Vaishampayan, and A. R. Reibman, "Redundancy rate-distortion analysis of multiple description coding using pairwise correlating transforms," in ICIP'97, 1997.
- [8] J. Kovacevic, "Multiple descriptions as joint source channel codes," 1998.
- [9] G.H. Golub and C.F. Van Loan, Matrix Computations, John Hopkins University Press, 1996.
- [10] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," Tech. Rep., Bell Laboratories, Lucent Technologies, 1996.
- [11] V. Hardman M. A. Sasse, M. Handley, and A. Watson, "Reliable audio for use over the internet," in Proc. INET, 1995.
- [12] Anil K. Jain, Fundamentals of Digital Image Processing, Prentice Hall Information and System Sciences Series. Prentice Hall, Englewood Cliffs, NJ 07632, 1989.
- [13] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, Numerical Recipes in C, Cambridge University Press, 2nd edition, 1992.
- [14] W. Jiang and A. Ortega, "Karhunen Loeve Vector Transform (KLVT) for vector quantization," Tech. Rep., University of Southern California, Los Angeles, CA, USA, 1996.