On Joint Distribution Modeling in Distributed Video Coding Systems

Yevgeny Priziment and David Malah

Department of Electrical Engineering,
Technion IIT, Haifa 32000, Israel
evgeniyp@tx.technion.ac.il, malah@ee.technion.ac.il

***Abstract*—Performance of a distributed video coding system depends, to a large extent, on the accuracy of joint source and side information distribution modeling. In this work we first examine a family of stationary joint distribution models. As one of our findings, we propose to use the double-Gamma model as an alternative to the widely adopted Laplace model, due to its superior performance. In addition, we suggest a new spatially adaptive model, which enables to follow the spatially varying joint statistics of the source and side information. We present two methods, class-based and neighborhood-based, for estimation of the spatially varying model parameters. We then show how the obtained pixel domain model can be used in the transform domain to facilitate utilization of frame spatial redundancy. Integration of the proposed models into a distributed video coding system resulted in improved performance.**

I. Introduction

Distributed Video Coding (DVC) is a novel coding scheme that employs principles of lossy source coding with side information (SI) at the decoder, also known as Wyner-Ziv (WZ) coding [1], [2]. In a baseline DVC system [1], the video sequence is split into Key frames and WZ frames. The Key frames are encoded by standard intra coding techniques while the WZ frames are encoded using syndromes (channel code) and decoded by combining these syndromes with the SI. In a DVC system, the SI is a prediction of the WZ frame generated at the decoder by motion compensated interpolation or extrapolation of the previously decoded Key and/or WZ frames.

The distributed video coding framework enables to shift the computational load of motion estimation from the encoder to the decoder, resulting in reversed encoder-decoder complexity. This reversed complexity scheme could be appealing for applications in which the encoder is power and/or complexity constrained, such as wireless/cellular video, sensor networks and video surveillance. The operation of a DVC system is based on a joint distribution model of the source and side information.

The overall performance of a DVC system is highly affected by the accuracy of the joint distribution model. The joint distribution is usually modeled by some distribution family with a set of adjustable parameters [3]. These parameters can be preset according to prior information, like off-line modeling or alternatively, they can be adapted online, according to the varying statistics. In this work we study the modeling problem in detail and analyze the performance of several joint univariate and multivariate distribution models, which were not considered in the context of DVC earlier.

MMSP'10, October 4-6, 2010, Saint-Malo, France.

This work is organized as follows. In Section II, we examine a set of models in an off-line mode in order to obtain access to both the source and the side information. We then integrate the most suitable model into the DVC system. Next, in Section III, we present a spatially adaptive model along with a method to estimate its parameters in the pixel domain, and show how to utilize these parameters in the DCT domain. Simulation results are given in each section. This work is concluded in Section IV.

II. Stationary Modeling of Joint Distribution

The Laplace distribution is a very popular model for the virtual channel in both the pixel and the transform domains [3], [4]. It is widely adopted due to its simplicity and good accuracy. In some works, it is used to model the transition probabilities of the virtual channel, i.e., the distribution of the side information given the source, while in others it is used to model the posterior distribution, i.e., the distribution of the source given the side information.

In [5], alternative models for the virtual channel were considered. The Laplace model was compared to Generalized Gaussian (GG), Gaussian and double-Gamma models. The robustness of these models was compared with respect to a mismatch in a model parameter (virtual channel noise variance) estimation. The double-Gamma and the GG models were found to be more robust, i.e., less sensitive to the variance estimation error. However, the simulations were performed in the *pixel domain* only. Moreover, the conclusions were drawn based on two frames only. In [6], inspired by models for traditional predictive video coding systems, a Wyner-Ziv rate-distortion function for a first order Markov-Laplace model was analyzed. Nevertheless, the model was tested only on synthetic signals and was not tested in a real DVC system.

In our study we also use the Laplace distribution as a reference model to which we compare the set of our test models. Estimation of the Laplace distribution scale parameter, α, can be interpreted as matching the variance of the model, $\sigma^2 = 2/\alpha^2$, to that of the empirical data. However, when examining higher order characteristics of the Laplace distribution, such as *kurtosis* $\kappa = \mu_4/\sigma^4$ (the peakedness of the distribution, where μ_4 and σ are the forth central moment and the standard deviation, respectively), it rarely matches the one obtained empirically. This is mainly because the kurtosis of the Laplace distribution is independent of α and equals to 6.

In order to allow for more accurate and flexible modeling, we study the performance of models in the Generalized Gamma Distribution (GΓD) [7] family, which is given by:

$$f_X(x; a, v, p) = \frac{pa^{-pv}}{2\Gamma(v)} |x|^{pv-1} e^{-\left(\frac{|x|}{a}\right)^p}, \quad (1)$$

for $x \in \mathbb{R}$ and $a, v, p > 0$, where $\Gamma(z) = \int_0^\infty x^{z-1} e{-}x dx$ is the Gamma function. GΓD is highly versatile and can be degenerated into various other distributions. For instance, the Generalized Gaussian distribution is obtained for $vp = 1$, the Normal distribution for $p = 2$ and $v = 0.5$, the Laplace distribution for $p = 1$ and $v = 1$, and a degenerate double-Gamma distribution with shape parameter of 0.5 is obtained by setting $p = 1$ and $v = 0.5$.

Parameters estimation of the GΓD has been an active research topic for several decades. Unfortunately, maximum likelihood (ML) estimation leads to transcendental equations that can only be solved graphically or numerically [7] and [8]. Maximum likelihood estimators for the GGD, double-Gamma, Laplace and Normal distributions can be obtained by substituting the specific values of a, v and p, corresponding to these distributions, into the maximum likelihood estimation equations for GΓD.

In order to compare the different models we evaluate their goodness of fit to the *virtual correlation channel* of the Distributed Video Coding system. The test is performed in an off-line mode and both Wyner–Ziv frames, X, and side information frames, Y, are assumed to be available. In addition, it is assumed, as accustomed in DVC systems [3], that the noise samples can be characterized as additive, white and independent of the side information Y (or alternatively of the WZ frame X). A DCT is applied to the difference $X - Y$ and the distribution parameters are estimated for each of the DCT coefficient bands. The goodness of fit is measured by two metrics: the Akike Information Criterion (AIC) [9] and the Minimum Description Length (MDL) [10]. These metrics are defined as follows:

$$AIC(\hat{\theta}_{ML}) = -2\log\left[f\left(\underline{x}; \hat{\theta}_{ML}\right)\right] + 2k, \quad (2)$$

$$MDL(\hat{\theta}_{ML}) = -\log\left[f\left(\underline{x}; \hat{\theta}_{ML}\right)\right] + \frac{k}{2\log(n)}, \quad (3)$$

where θ_{ML} is the maximum likelihood parameters set, n is the sample population size and k is the number of model parameters. As can be inferred from (2) and (3), a model with a better fit to the data will produce smaller MDL and AIC values. Note that both metrics penalize distributions with large number of parameters in order to avoid over-fitting.

We performed simulations on standard test sequences, listed in Table I, all at CIF resolution. For each sequence, a Group of Pictures (GOP) size of 2 was used and the side information frame was obtained using motion compensated interpolation [11]. The motion was estimated from two Key frames, adjacent to the WZ frame, using full-search block matching. Maximum likelihood parameters were estimated for the following distributions: Generalized Gamma, Generalized Gaussian, double-Gamma, Laplace and Gaussian. Next, MDL and AIC metrics were calculated. The results were averaged over all frames and across the DCT bands and are summarized in Table I.

TABLE I
MDL AND AIC AS GOODNESS OF FIT METRICS FOR DCT DOMAIN VIRTUAL CHANNEL

Sequence		GΓD	dbl.-Gamma	GGD	Laplace	Gauss.
Mobile	AIC	**1.80e+04**	**1.80e+04**	1.92e+04	2.00e+04	2.25e+04
	MDL	**9.01e+03**	9.03e+03	9.62e+03	9.99e+03	1.13e+04
Coastguard	AIC	**1.64e+04**	1.65e+04	1.73e+04	1.74e+04	1.87e+04
	MDL	**8.22e+03**	8.25e+03	8.66e+03	8.70e+03	9.37e+03
Container	AIC	**7.98e+03**	8.23e+03	1.09e+04	1.10e+04	1.24e+04
	MDL	**4.00e+03**	4.12e+03	5.45e+03	5.49e+03	6.22e+03
Foreman	AIC	**1.48e+04**	1.49e+04	1.64e+04	1.67e+04	1.82e+04
	MDL	**7.41e+03**	7.45e+03	8.22e+03	8.34e+03	9.11e+03
Flower Garden	AIC	**1.41e+04**	**1.41e+04**	1.72e+04	1.93e+04	2.19e+04
	MDL	**7.04e+03**	7.06e+03	8.61e+03	9.65e+03	1.10e+04
Hall Monitor	AIC	**1.07e+04**	1.10e+04	1.44e+04	1.44e+04	1.55e+04
	MDL	**5.33e+03**	5.50e+03	7.19e+03	7.21e+03	7.74e+03

Analyzing the results, one can see that GΓD outperforms all other examined models, for all tested sequences, both for AIC and MDL metrics. However, the performance of the double-Gamma distribution is very close to that of GΓD and for some sequences even matches it. This observation is very inspiring since double-Gamma distribution has only one parameter, as opposed to 3 for GΓD and 2 for GGD. Moreover, the double-Gamma distribution parameter can be obtained using a closed form ML estimator given by:

$$\hat{a}_{ML} = \frac{2}{n}\sum_{i=1}^{n} |x_i|, \quad (4)$$

where $\{x_i\}_{i=1}^n$ are the n samples used for estimation.

It should be noted that the sample population size implicitly affects both MDL and AIC metrics through the ML (first) term of (2) and (3). For CIF resolution, assuming that a whole frame is a single slice, each DCT coefficient band has 6336 elements which makes the penalizing terms of (2) and (3) negligible and results in a factor of 2 difference between AIC and MDL. Aiming to keep the modeling and the subsequent estimation problems as simple as possible and based on these off-line observations, we selected the double-Gamma model as a candidate to replace the traditional Laplace model.

We compared the performance of the double-Gamma based model vs. the Laplace based model for *Coastguard* at QCIF resolution and *Flower Garden* at CIF resolution test sequences. The utilized DVC system is based on the design proposed by the Stanford group [1]. The Key frames were intra coded using H.264. The WZ frames were syndrome coded using LDPC codes, applied to quantization indices of 4×4 DCT coefficients. The quantization steps were adjusted to obtain constant quality. At the decoder, the side information was generated using motion compensated interpolation of the Key

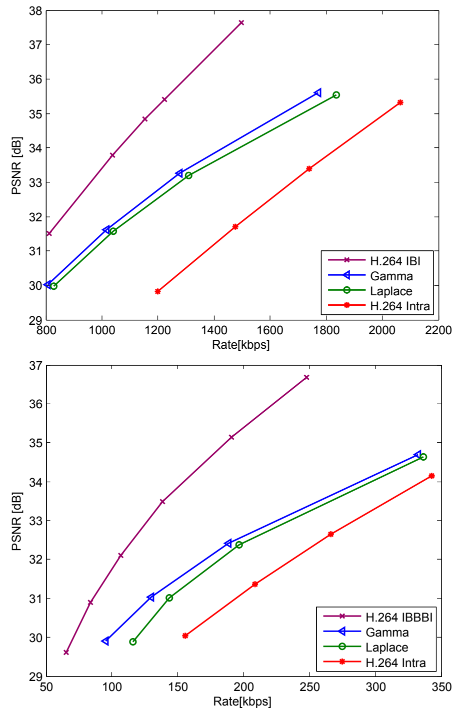

Fig. 1. Comparison of virtual channel models for *Flower Garden* (**top**) and *Coastguard* (**bottom**) sequences.

frames [11]. The *Coastguard* sequence is coded with a GOP size of 4 while the *Flower Garden* with a GOP size of 2. The Rate-PSNR curves obtained for these sequences are depicted in Fig. 1. The bit rate corresponds to a frame rate of 15 frames per second. In addition, the performance of the DVC system is compared to that of a baseline H.264 operating in Intra mode, denoted by 'H.264 Intra', or in Inter mode denoted by 'H.264 IBI' for a GOP size of 2 and similarly by 'H.264 IBBBI' for a GOP size of 4. Comparing the rates, it can be seen that adopting the double-Gamma model instead of the Laplace model results in a decrease in the coding rate by at least 2.5% and up to 4% for *Flower Garden*, and by 2% to 18% for *Coastguard*.

III. Spatially Adaptive Modeling

The modeling approach discussed in the previous section relies on the assumption that the virtual channel is stationary across the whole frame or across a number of slices tiling the frame. In practice, this means that a model with a single set of parameters is used for all pixels, in the pixel domain DVC system, or for every DCT coefficients band in the transform domain system. The stationarity assumption is obviously an oversimplification of the actual properties of natural video sequences. The correlation characteristics vary across the frame as a function of the prediction (side information) accuracy at each pixel. Side information accuracy is affected by various factors such as image spatial content (edges, texture), camera and objects motion and algorithms used at the decoder for motion estimation and compensation. Accurate model adaptation to the spatially varying statistics has the potential to significantly improve the coding efficiency of a DVC system.

Spatially adaptive modeling in a pixel domain DVC system was first studied by Brites et.al., in [11] and [3]. They proposed to estimate the Laplace model parameters at frame, block and pixel levels. Peculiarly, for certain combinations of frame level and block level variances, the pixel level model parameter is estimated from a single sample, while for other scenarios the block or frame level parameter is adopted. Nevertheless, rate reductions of up to 5% were reported for pixel level relatively to the frame level modeling. A similar approach was adopted by Brites et. al., for the transform domain modeling and a rate decrease of 3%, on average, was reported for coefficient level relatively to a band level modeling. First, a frame level parameter is estimated for each DCT band. Second, for each coefficient with squared magnitude, m^2, larger than the band variance, the model parameter is set to $\sqrt{2}/m$. A more interesting approach was presented in [4] by Skorupa et. al. First, it was proposed to estimate a block level variance in the pixel domain. Next, a set of transform domain coefficients variances was obtained by considering each DCT coefficient as a linear combination of the pixel domain components. However, the resulting transform domain covariance model relies on some empirically crafted expressions. Moreover, it has a fixed form which depends only on the pixel domain block variance.

Here we follow a similar approach to that discussed in [4]. However, we allow for different forms of the autocovariance function and show how the transform domain parameters can be analytically obtained from those in the pixel domain without any approximations. The virtual channel is modeled by a MultiVariate Laplace (MVL) distribution over blocks of $n_b \times n_b$ pixels in the spatial domain. The MVL distribution was selected due to its invariance to linear transformations, a property that will shortly prove to be useful. It is worth mentioning that this property could not be verified for the multivariate double-Gamma distribution.

The MVL distribution is a generalization of the univariate Laplace distribution and is defined as follows [12]:

$$f(z) = 2(2\pi)^{-\frac{d}{2}} |\Sigma|^{-\frac{1}{2}} \left(\frac{z'\Sigma^{-1}z}{2}\right)^{\frac{v}{2}} K_v\left(\sqrt{2z'\Sigma^{-1}z}\right), \quad (5)$$

for $z \in \mathbb{R}^d$, where Σ is a positive definite autocovariance matrix, $d = n_b^2$, $v = (2-d)/2$ and $K_v(\cdot)$ is the modified Bessel function [13]. In the following, we refer to the elements of the autocovariance matrix as σ_{ij}, $0 \leq i, j \leq d$.

Alternatively, the MVL distribution can be defined by its characteristic function [12]:

$$\Phi(t) = \frac{1}{1 + \frac{1}{2}t'\Sigma t}, \quad t \in \mathbb{R}^d. \quad (6)$$

Marginalization of the MVL distribution results in lower dimension MVL distribution or in the familiar univariate Laplace distribution.

In our virtual channel model, the MVL autocovariance matrix, Σ, varies from block to block, enabling to follow the spatial variations. The elements of the autocovariance matrix are estimated in the pixel domain from the noise image, $Z = X - Y$. However, statistical characterization of transform domain virtual channel is needed in order to facilitate utilization of the spatial correlation.

Similarly to the Multivariate Gaussian distribution, linear transformation, A, of a Multivariate Laplace distributed vector $X \sim MVL(\Sigma)$, results in a Multivariate Laplace distributed vector $Y \sim MVL(A\Sigma A')$. This property can be easily verified by examining the characteristic function of $Y = AX$, as following:

$$\Phi_Y(t) = E[e^{iY't}] = E[e^{i(AX)'t}] = \Phi_X(A't). \qquad (7)$$

Consequently, since the DCT is a linear transform, the spatial characteristics of the virtual channel can be transformed to obtain spatially varying characterization of the noise image DCT coefficients. Next, we present two methods for pixel domain MVL parameters estimation.

1) Neighborhood-based Virtual Channel Modeling: In order to estimate the MVL parameters, a decoder-side noise image is constructed, following the steps in [14]. First, motion estimation is carried out between two reference frames, one preceding and the other following the Wyner-Ziv frame to be decoded. Second, motion compensation to the time instant corresponding to the Wyner-Ziv frame is performed on these reference frames. The average of these compensated versions is used as the side information, Y. However, since X is not available at the decoder, the true noise image $Z = Y - X$ can not be obtained. Instead, the decoder constructs a noise image approximation as half of the difference between the compensated reference frames [3].

Once the noise image is at hand, we propose to estimate a local autocovariance matrix for a each block by examining the block's close neighborhood within a window $x_w(\cdot,\cdot)$ of $n_w \times n_w$ pixels. The estimate of the autocovariance matrix is obtained using the standard sample autocovariance estimator as follows:

$$Cov(u,v) = \frac{1}{(n_w - u)(n_w - v)} \times \sum_{\substack{|k-i|=u,\\ |l-j|=v}} x_w(k,l)x_w(i,j). \qquad (8)$$

It should be noted that this approach might suffer from instability, i.e., the estimated autocovariance matrix might not be positive definite. On the one hand, we would prefer the neighborhood window to be as small as possible in order to obtain better localization of the model. On the other hand, for a small sample size, the estimated parameters are less reliable. Moreover, the estimated autocovariance matrix might be non-positive definite, implying that the distribution is degenerate. A possible solution to these problems is achieved by adopting a model for the autocovariance matrix, such as a first order exponential model given by $Cov(u,v) = \sigma^2 \exp(\rho_1 u^2 + \rho_2 v^2)$ and estimating the model parameters, in this case ρ_1 and ρ_2. Another option is to utilize shrinkage estimation methods [15]. We adopt a simpler approach, first we transform the autocovariance matrix to the DCT domain, as described bellow, and then replace all elements on the main diagonal that are below some predefined threshold, $thr > 0$, by one half of the smallest legitimate variance: $\min\{\sigma_{ii} | \sigma_{ii} > thr\}$.

The DCT domain distribution parameters are obtained by transforming the covariance matrix according to (7) with A given by the 2D DCT transform. The obtained transform domain autocovariance matrix contains the DCT coefficients' variances across the main diagonal, while other matrix elements are the cross terms. Typically the cross terms have very small values, relative to the variances due to the de-correlating property of the DCT transform. Two examples of DCT domain autocovariance matrices are presented in Fig. 2. Consequently, the cross-band correlation can be discarded without significant loss in the system's performance. As a result, each DCT coefficient is assumed to be a sample of a univariate Laplace Distribution with a scale parameter given by $\alpha_i = \sqrt{2/\sigma_{ii}^2}$, where σ_{ii}^2 is the DCT coefficient variance obtained from the main diagonal of the transform domain autocovariance matrix. It should be noted, that though each DCT coefficient is modeled by a univariate Laplace distribution, each coefficient has a different scale parameter which is explicitly affected by the spatial variations in the noise image. This is in contrast to stationary modeling, where DCT coefficients within each band are modeled by univariate Laplace distribution, which has a single fixed scale parameter α across the whole frame.

2) Class-based Virtual Channel Modeling: The reliability of the estimation process can be improved by increasing the sample size. Nevertheless, simply enlarging the neighborhood window, $n_w \times n_w$ will result in a loss of locality. Alternatively, the sample size can be increased by grouping blocks with similar characteristics, like foreground or background, static or dynamic, textured or smooth and their combinations. However, classifying the blocks into various classes is a complex task. Here, we consider a simplified classification approach which uses only two classes as described below.

When examining an image of the virtual channel noise (see Fig. 3 top) together with the Wyner-Ziv frame, it is easy to see that the blocks of the noise image can be roughly classified into two sets. One set containing the blocks with the scene background and the other containing the blocks with moving objects' edges, we denote these sets by S_{bkg} and S_{edg}, respectively. Typically, the scene background is static or undergoes global motion and thus can be predicted quite accurately by the decoder, resulting in relatively low noise. On the other hand, around moving objects, some pixels are revealed while others are occluded. This makes the prediction around the edges less accurate, which leads to higher noise.

We propose to estimate the MVL parameters for each class separately. First, we use one of the compensated reference frames to locate the edges in the image, applying the Sobel edge detector [16], see Fig. 3 bottom. The obtained edges are dilated using morphological operators, in order to imitate

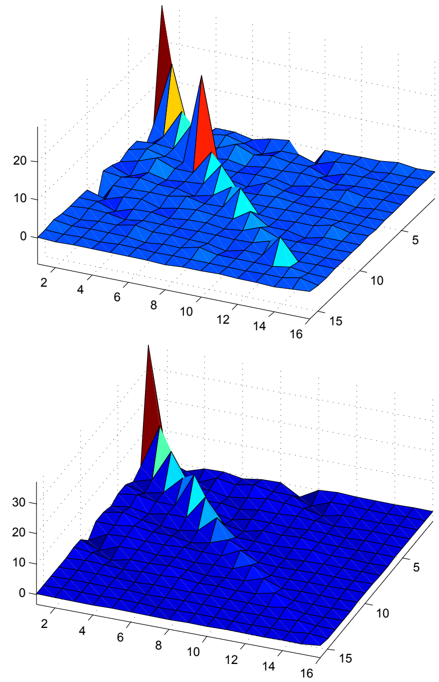

Fig. 2. Sample autocovariance in the DCT domain, obtained by transforming the pixel domain autocovariance matrices using the Neighborhood based approach with window size $n_w = 8$ and a block size of 4×4. In DCT domain the energy is concentrated on the main diagonal.

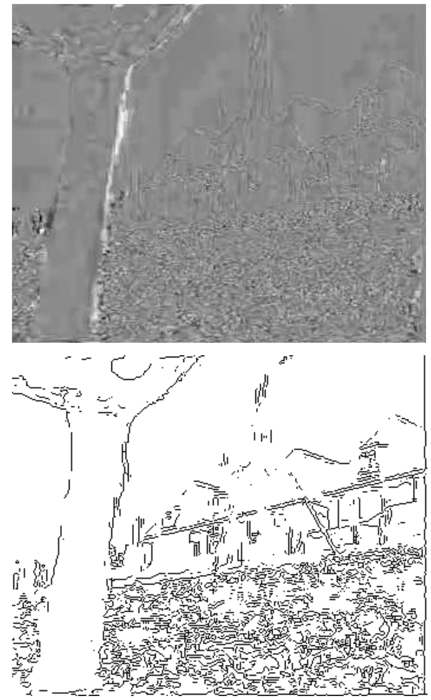

Fig. 3. Approximated decoder side noise image (**top**) and the corresponding edge image (**bottom**) obtained by applying Sobel edge detector to motion compensated preceding reference frame for the 244th frame of *Flower Garden* sequence. Higher error values are obtained along the edges in the image.

the smearing effect around the edges due to the motion. The structuring element used in the dilation process can be adapted to the sequence and frame characteristics. For instance, in fast motion sequences or when working with large GOPs (resulting in a large number of WZ frames between any two key frames), a larger structuring element is applied.

Next, the resulting edge image is used to classify the blocks into the two classes of scene background and moving objects, as mentioned earlier. This is performed in the following manner. The edge image is divided into $n_b \times n_b$ blocks. A soft classification method is then applied to each block, i.e., two weights, w_{bkg} and w_{edg}, are assigned to each block. The sum of these weights for each block equals 1 and the weight of the moving objects class is proportional to the number of pixels containing an edge in the block.

Similarly to the Neighborhood-based approach, model parameters are estimated using the decoder's version of the noise image. In each class, the blocks, b_m, of the noise image are considered as samples of an $n_b \times n_b$ multivariate Laplace distribution and are used for sample autocovariance matrix estimation.

$$Cov_{class}(u,v) = \frac{1}{(n_b-u)(n_b-v)\sum_{b_m \in S_{class}} w_{m,class}} \times \sum_{b_m \in S_{class}} \sum_{\substack{|k-i|=u,\\ |l-j|=v}} w_{m,class} \cdot x_{b_m}(k,l) \cdot x_{b_m}(i,j) \quad (9)$$

The estimated autocovariance matrices are associated with the image blocks according to the original classification. Blocks with mixed content (containing both background and objects' edges) are modeled as a mixture of the two MVL distributions.

Once the autocovariance matrix is estimated for each block, we proceed as before. The transform domain statistics are obtained by block-wise transformation of the autocovariance matrix estimates. Similarly to the Neighborhood-based method, the energy is concentrated on the main diagonal, see Fig. 4. Consequently, only the variance values are used, while the cross band correlations are discarded.

The performance of the proposed spatial adaptive MVL based system was compared to that of stationary Laplace and double-Gamma based systems. The results are depicted in Fig. 5, where the curve corresponding to the proposed Neighborhood-based method for estimation of the MVL parameters is denoted by 'MVL nbhd.' and the one corresponding to the Class-based method is denoted by 'MVL cls.'. As it can be seen, the MVL based system with neighborhood-based parameter estimation outperforms the double-Gamma, Laplace and the class-based MVL systems. A rate reduction of 6% to 10% can be observed for *Flower Garden* and 6% to 19% for *Coastguard*. Alternatively, gains of up to 1dB for *Flower Garden* and of 0.5dB for *Coastguard* are noticed.

IV. Conclusion

In this work we have considered the problem of source and side information joint distribution modeling. We have addressed both the model selection problem as well as the invalid spatial stationarity assumption. We have demonstrated that the Laplace model falls behind the Generalized Gamma

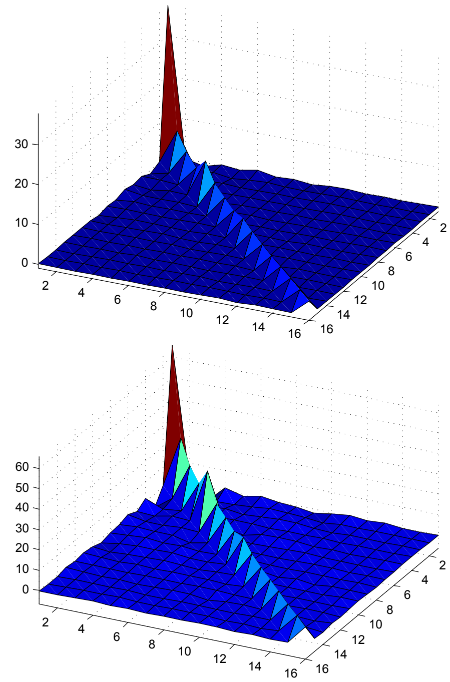

Fig. 4. Transform domain autocovariance matrix of the background class blocks, Σ_{bkg}, (**top**); and the edges class blocks, Σ_{edg}, (**bottom**). Obtained for the second frame of the *Flower Garden* sequence.

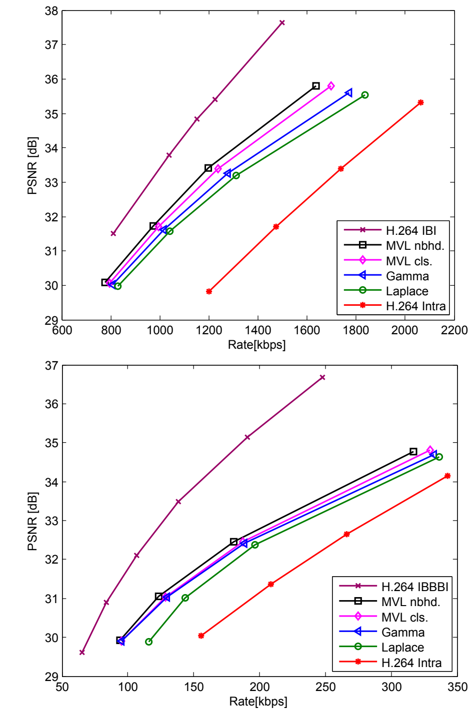

Fig. 5. MVL performance for *Flower Garden* (**top**) and *Coastguard* (**bottom**) sequences compared to the stationary Laplace and Gamma models and to H.264.

and the double-Gamma distributions. In addition, we have presented spatially adaptive modeling based on the multivariate Laplace distribution. Simulations of a DVC system based on the spatially adaptive model have shown that it outperforms both double-Gamma and Laplace stationary models.

There are several possible directions for future work, as we see it. First, further steps can be taken in modeling the spatially varying joined distribution. It is expected that more accurate information can be obtained by utilizing knowledge of the frame structure, such as edges and texture, along with a reliability metric on the estimated motion field. A second direction might be generalizing the spatially adaptive modeling into spatio-temporal adaptive modeling. Instead of estimating model parameters based on a window within a single frame, the estimation can be extended to several frames, such that it is performed within a 3D volume. The 3D volume should be aligned along the motion trajectory of the object.

Acknowledgment

This work was supported in part by a grant from Elbit Systems Ltd. The help of the staff of the Technion Signal and Image Processing Lab (SIPL) is greatly appreciated.

References

[1] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proc. IEEE*, vol. 93, no. 1, pp. 71–83, January 2005.

[2] F. Dufaux, W. Gao, S. Tubaro, and A. Vetro, "Distributed Video Coding: Trends and Perspectives," *EURASIP Journal on Image and Video Processing*, 2009.

[3] C. Brites and F. Pereira, "Correlation noise modeling for efficient pixel and transform domain Wyner-Ziv video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 9, pp. 1177–1190, Sept. 2008.

[4] J. Skorupa, J. Slowack, S. Mys, P. Lambert, and R. Van de Walle, "Accurate Correlation Modeling for Transform-Domain Wyner-Ziv Video Coding," in *Proc. Pacific Rim Conference on Multimedia: Advances in Multimedia Information*. Springer-Verlag, 2008, p. 10.

[5] R. Westerlaken, S. Borchert, R. Gunnewiek, and R. Lagendijk, "Dependency channel modeling for a LDPC-based Wyner-Ziv video compression scheme," in *Proc. IEEE ICIP*, 2006, pp. 277–280.

[6] V. Sheinin, A. Jagmohan, and D. He, "Uniform threshold scalar quantizer performance in Wyner-Ziv coding with memoryless, additive laplacian correlation channel," in *Proc. IEEE ICASSP*, vol. 4, May 2006.

[7] E. Stacy and G. Mihram, "Parameter estimation for a generalized gamma distribution," *Technometrics*, vol. 7, no. 3, pp. 349–358, 1965.

[8] J. Chang, J. Shin, N. Kim, and S. Mitra, "Image probability distribution based on generalized gamma function," *IEEE Signal Process. Lett.*, vol. 12, no. 4, pp. 325–328, 2005.

[9] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Proc. IEEE ISIT*, pp. 267–281, 1973.

[10] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.

[11] C. Brites, J. Ascenso, and F. Pereira, "Studying temporal correlation noise modeling for pixel based Wyner-Ziv video coding," in *Proc. IEEE ICIP*, 2006, pp. 273–6.

[12] S. Kotz, T. Kozubowski, and K. Podgorski, "An asymmetric multivariate Laplace distribution," *Mathematical and computer modeling*, vol. 34, pp. 1003–1021, 2003.

[13] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. National Bureau of Standards, 1972.

[14] C. Brites, J. Ascenso, and F. Pereira, "Modeling correlation noise statistics at decoder for pixel based Wyner-Ziv video coding," *Proc. PCS*, 2006.

[15] J. Schafer and K. Strimmer, "A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics," *Statistical applications in genetics and molecular biology*, vol. 4, no. 1, p. 1175, 2005.

[16] R. Gonzalez and R. Woods, *R. Digital Image Processing*. New Jersey: Prentice Hall, 2002.