A Feature Oriented Framework and Enhanced Assessments for Imaging Compression*

1st Jesse Redford
Department of Mechanical Engineering
University of North Carolina Charlotte
Charlotte, NC, USA
tredford@uncc.edu

2nd Xingjie Li
Department of Mathematics and Statistics
University of North Carolina Charlotte
Charlotte, NC, USA
xli47@uncc.edu

2020 IEEE 4th International Conference on Image Processing, Applications and Systems (IPAS) | 978-1-7281-7574-4/20/$31.00 ©2020 IEEE | DOI: 10.1109/IPAS50080.2020.9334966

***Abstract*—This paper offers a new feature-oriented compression algorithm for flexible reduction of data redundancy commonly found in images and videos streams. Using a combination of image segmentation and face detection techniques as a preprocessing step, we derive a compression framework to adaptively treat 'feature' and 'ground' while balancing the total compression and quality of 'feature' regions. We demonstrate the utility of a feature compliant compression algorithm (FC-SVD), a revised peak signal-to-noise ratio PSNR assessment, and a relative quality ratio to control artificial distortion. The goal of this investigation is to provide new contributions to image and video processing research via multi-scale resolution and the block-based adaptive singular value decomposition.**

***Index Terms*—adaptive singular value decomposition, feature oriented image compression, modified PSNR.**

I. Introduction

Recent studies have suggested that image compression improves image classification accuracy, for convolution networks such as Inception-3, trained on ImageNet database [1]. In addition, the growing number of mobile computing platforms and applications which rely on image processing systems has made the transmission and storage of digital image data increasingly challenging. Digital image compression is a technique which is used to create more compact representations of available data without compromising essential associations and details. In this paper, we propose a feature-oriented compression approach and associated assessments which extends Region of Interest Based Coding techniques to more efficiently compress data for storage and/or transmission.

In the domains of computer vision and image processing, a "feature" is often referred to as a piece of relevant information that is required for solving a computational task associated with an application. Although there is a variety of general image and video compression algorithms available, such as JPEG or MPEG standards, these types of compression processes often neglect the presence of particular features found in emerging image processing applications, such as face recognition. Hence, they are not implicitly designed for objective/feature oriented compression. For this reason the utility of these compression standards is often limited to general purpose compression applications. Similarly, the standard peak-singal-to-noise-ratio ($PSNR$) assessment which provides an objective measure of the reconstruction quality of compressed images, does not provide sufficient evaluations to best analyze the performance of a feature oriented image compression scheme (see e.g., [2]–[5]). Thus, there is a need for reliable objective assessments which provide approximate measures for feature quality and convey relative distortion.

This project is funded by the NSF CAREER DMS-1847770.

The use of singular-value-decomposition (SVD) to create more compact representations of image data by truncating the singular components has been widely studied [6]. For SVD compression an emphasis is placed on the determination of the amount of truncation which balances the level of compression and the degradation of the image. To quantify the amount of degrading, an appropriate objective metric is needed to measure the quality of the resulting image and efficiency of compression.

Although, it is often purpose specific and domain related, for image compression selecting an objective metric which agrees with Human Visual Perception or region of interests is often advantageous. In this study, we employ adaptive block-based implementations and develop new assessment metrics to enhance the overall compression quality. Using a combination of image segmentation and feature detection techniques as a pre-processing step, we draw upon developed visual system methodologies to derive a feature-oriented compression criterion for controlling the local components of SVD truncation.

Outline of the paper. Section II includes an overview of SVD image compression, $PSNR$ quality assessment, and details an image partitioning scheme for enhancing adaptive block-based implementations. In section III we discuss a formulation of a revised $PSNR$ assessment for image quality and define an additional distortion measure Q to quantify and control the artificial distortion. Section IV includes the methodology and algorithmic details related to our construction of a feature oriented compression framework. Section IV also includes the underlying reasoning of the proposed metrics. Examples are provided to demonstrate the utility of these proposed assessments in the evaluation of compression schemes. Furthermore, Section V evaluates the performance of the proposed FC-SVD algorithm in comparison to other compression methods in still image and dynamic video processing applications. Section VI summarizes the paper.

II. Singular Value Decomposition

A. Review of SVD

Singular value decomposition and truncation of the singular components are commonly used techniques to achieve image compression and processing. Given an image with $m \times n$ pixels, representing this image as matrix $I_{m \times n}$, requires $m \times n$ memory locations. Leveraging the global SVD decomposition to approximate the matrix I by using only k singular value components, yields an alternative representation requiring only $k(m+n+1)$ memory locations to store the image. The following condition describes the constraint on the rank approximation-k required to achieve compression.

$$CF(k) = \frac{m \times n}{k\,(1+m+n)} \quad \geq 1. \tag{1}$$

B. Block-Based Implementations

Common SVD compression schemes often entail local and block-based implementations, which often yields better compression performance and reduces processing time. For SVD performance factors such as image reconstruction quality and compression rates depend on the estimated rank used to represent a given image matrix, hence in a block-based implementation the compression performance is contingent upon the series of rank estimates used to describe the approximation of the image matrix as a whole. For cases in which a series of sub-matrix decompositions are used to represent the image matrix. Given image matrix I with dimensions $M \times N$, and an area which is divisible by a collection of smaller matrices of equal dimension $m_{b_i} \times n_{b_i}$, for each sub-block of the image matrix, one can truncate its singular components to k_i differently by incorporating various regions of interests.

C. Adaptive Partitions

Because features within an image can vary in shape and size, given an asymmetrical case described by Figure 1, where an image is partitioned into multiple sections.

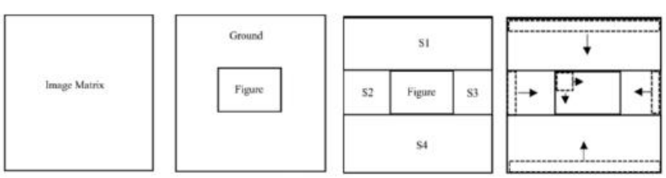


Fig. 1: Visual representation of matrix segmentation for feature and non-feature partitions

We introduce Equation (2) in order to quantify the compression factor of multiple sub-matrices.

$$\widetilde{CF}^{b} = \frac{MN}{\sum_{s=1}^{S_n}\left[\sum_i k_i\,(1+m_{b_i}+m_{b_i}) \times \left(\frac{M_s}{m_{b_i}}\frac{N_s}{n_{b_i}}\right)\right]}, \tag{2}$$

where M, N are the dimensions of the entire image matrix, and S is a partition of multiple sections with dimensions $M_s \times N_s$; each section is divided by a few smaller matrices with dimensions $m_{b_i} \times n_{b_i}$. These extensions of described partition cases are demonstrated as the one given in Figure 1.

D. Review of Objective Measures for Image Quality

Conventional practice for image processing such as rate-distortion studies on both the compression of still images and video sequences use $PSNR$ as the base visual quality metric [7]. The objective assessment is an approximate estimation to human perception of the reconstruction quality of compressed image [8]. Given the original image I, the $PSNR$ for the compressed one I_k is defined as

$$\begin{aligned} &PSNR(I) \\ &= 10 \cdot \log_{10}\left(\frac{\mathrm{MAX_I^2}}{\frac{1}{\mathrm{mn}}\sum_{\mathrm{i=0}}^{\mathrm{m-1}}\sum_{\mathrm{j=0}}^{\mathrm{n-1}}\left[\mathrm{I\,(i,j)} - \mathrm{I_k\,(i,j)}\right]^2}\right) \end{aligned} \tag{3}$$

where m and n are the number of rows and columns of the image matrix and $\mathrm{MAX_I}$ is the maximum possible pixel value, calculated as $2^b - 1$, where b represents the number of bits used to represent the image. $PSNR$ is often expressed on the logarithmic decibel scale, denoted as dB. Typical $PSNR$ values for lossy image compression range between 30 and 50dB for 8 bit images, 60 to 80dB for 16 bit images [9], [10], with wireless transmission quality often accepting a range between 20 and 25db [11]–[13]. Note that $PSNR$ shares a positive correlation with image quality, so a higher $PSNR$ indicates a higher quality image.

III. Feature-Oriented Assessments

Although widely used, $PSNR$ does not account for pixel changes in feature locations and their perceptual importance in the quality evaluation of the observable scene [7]. In the proposed implementation, we would like to divide the entire image I into 'feature' part F and 'ground' part G such that the given image of size $M \times N$ is divided into F non-overlapping blocks of 'features' and one remaining portion of 'ground'. Each 'feature' block is indexed by s_i with size $M(s_i) \times N(s_i)$, so the total number of pixels of the 'features' is $P_F := \sum_{s_i=1}^{F} M(s_i) \times N(s_i)$, whereas the number of pixels within the 'ground' equals $P_G := M \times N - P_F$. We define the weighted mean $PSNR(F)$ to account for the contribution of each 'feature' object below

$$PSNR(F) = \sum_{S_i=1}^{F} w_{s_i} \cdot PSNR(s_i), w_{s_i} = \frac{M(s_i)N(s_i)}{P_F} \tag{4}$$

with $PSNR(s_i)$ denoting the $PSNR$ value restricted on the i-th 'feature' object $F(s_i)$:

$$\begin{aligned} &PSNR(s_i) \\ &= 10 \cdot \log_{10}\left(\frac{MAX_I^2}{\frac{1}{M(s_i)N(s_i)}\sum_{(i,j)\in F(s_i)}\big(I(i,j) - I_k(i,j)\big)^2}\right). \end{aligned}$$

We also define $PSNR(G)$ for the 'ground' portion as

$$PSNR(G) \tag{5}$$
$$=10 \cdot \log_{10}\left(\frac{MAX_I^2}{\frac{1}{P_G}\sum_{(i,j)\in G}\big(I(i,j)-I_k(i,j)\big)^2}\right).$$

Recall the definition (3) of $PSNR(I)$ for a global image and notice that $-\log(\cdot)$ is a convex function, so by the Jensen's inequality, we have

$$\frac{P_F}{MN}PSNR(F)+\frac{P_G}{MN}PSNR(G) \geq PSNR(I).$$

Because the discrepancy is always positive, guaranteed by the Jensen's equality, we show that the weighted sum of the $PSNR(G)$ and $PSNR(F)$ assessments yield a compatible measure for calculating $PSNR(I)$, that is Equation (6)

$$\widetilde{PSNR}(I) := \frac{P_F}{MN}PSNR(F)+\frac{P_G}{MN}PSNR(G). \tag{6}$$

For a compressed image with controllable distortion, we can assume that

$$\widetilde{PSNR}(I) \approx PSNR(I). \tag{7}$$

A. Measuring Relative Distortion

The proposed $PSNR(F)$, $PSNR(G)$ evaluations have the added benefit of obtaining error quantities representing the respective $PSNR$ for 'feature' and 'ground' regions separately. This also allows for the derivation of a relative distortion measure, denoted as $Q_{F,G}$, in Equation (8), which serves as a scaling percent measure of the distortion between the 'feature' and 'ground' spaces of a compressed image.

$$Q_{F,G} := \left(1-\frac{PSNR(G)}{PSNR(F)}\right)\times 100\% \tag{8}$$

Intuitively, this measure provides an effective assessment to evaluate the emphasis a compression process places on the preservation of the feature space in the image, with respect to the ground space. Notice $Q_{F,G} \sim 0$ % indicates homogeneous reconstruction error of the image, and the absence of emphasis placed on the compression of either ground or feature regions. Under this definition it is also assumed that the image being evaluated has compression of both 'feature' and 'ground' regions, such that neither $PSNR(F)$ or $PSNR(G)$ is equal to infinity.

IV. Feature-Oriented Compression

A. Overview of methodology

The perceived environment always presents far more perceptual information than the human-visual-system (HVS) can effectively process. As a result, certain regions within an image are more relevant to the HVS. In this section, we explore the integration of these concepts into a compression criteria by defining a region-of-interest(s) (ROI) within image frames. We refer to the image's ROI's as 'features', referencing the remaining portions of the image as 'ground'. This distinction is leveraged to tailor compression rates in the spatial domain by prioritizing computational attention to different regions of an image or video stream.

Compressing an image by automatically allocating more bits to region(s) of interest relative to an image's ground space has multiple advantages [14]. First, one can ensure vital information is protected. Second, is the reduction of substantial redundancy. Third, is enhanced compression efficiency for transmission and storage without compromising accuracy [14], [15]. In addition, our results indicate that image deterioration may be less obvious to the human eye when present in areas of an image which do not receive the same level of visual attention. Thus, the proposed methods may enhance compression for image types such as portrait, crowd, and image processing applications such as face recognition.

B. Algorithmic Implementation

The proposed feature compliant SVD compression scheme (FC-SVD), is given in Algorithm 1. The implementation is intended to be agnostic to a particular image segmentation process, but assumes a set of feature points $F(x_i, y_i)$ are available to construct boundary conditions, given by some detection model Θ. For our results we utilize OpenCV face detection model to provide these boundary conditions.

Algorithm 1 FC-SVD feature compliant SVD compression

Require: Matrix X_{in}, Detection model Θ, Low-rank scheme Ψ, Error threshold ξ, Desired Compression factor Φ

Ensure: Compressed Matrix X_{out}

```
X_comp = [0]_{M×N}
F(x_i, y_i) = f(Θ, X_in)
S[m_i, n_i] = (M|m', N|n') ∈ Φ
for s = m_i, n_i in S_i do
    for i in range(0, M, m_i) do
        for j in range(0, N, n_i) do
            X_b = X_in[i : i + m, j : j + n]
            X_b ← [U, Σ, V^T]_{X_b}
            Ψ = { Ψ_feature   X_b ∈ F(x_i, y_i)
                { Ψ_ground    X_b ∉ F(x_i, y_i)
            X_comp ← X_b^Ψ = Σ_{i=1}^{Ψ} σ_i u_i v_i^T
        end for i
    end for j
    if X_comp ∈ (ξ, Φ) then X_out = X_comp
    end if
end for s
return X_out
```

Although we do not offer a rigorous criteria for how one should optimize or select parameters values in Algorithm 1, we provide some empirical suggestions below. Also notice that in the above, we repeat the compression process using a series of divisible block sizes m_i, n_i stored in S_i until a desired performance criteria is met, hence an optimal block-size for a low-rank scheme is also determined.

C. Performance Analysis

Figure 2 provides examples of the proposed measures (2) (4) (5) (6) (8) on compressed images of a crowd processed by Algorithm 1 with empirical parameters. We compare the uniform block, the feature-oriented, and an example of ground-oriented compression, respectively. For each example, a block-size of 20x20 was used to compress the 8bit 1200x1920 image and the truncation of singular components for each example is treated differently.

Demonstrated by Figure 2 and the measured quantities, we recommend to keep the constraint $0 \leq Q_{F,G} \leq 20\%$ for feature- oriented compression as this avoids a noticeable compression contrast between 'feature' and 'ground' regions, though we don't have a rigorous reason behind this. Notice the $PSNR(F)$ of 33.24db is higher than the $PSNR(I)$, highlighting the advantage of measuring the quality of region of interest, rather than the entirety of the observable scene in an image.

$PSNR(F)$ yields a good objective assessment relative to $PSNR(I)$, based on a subjective visual measure of the detected crowd faces with respect to the ground area. In addition, the proposed metrics $\widetilde{PSNR}(I)$ yield a reliable approximation for a global $PSNR(I)$ assessment. Our proposed extensions allow for an independent and relative evaluation of the reconstruction quality in both the 'feature' and 'ground' regions. The results shown in Figure 2 indicate that our algorithm provides enhanced reconstruction quality, while maintaining comparable space savings to global and uniform block SVD. Notice the proposed algorithm FC-SVD achieves a CF of 7.76 with a $PSNR(F)$ of 36.44, whereas the uniform block compression achieves a CF of only 3.24, with the same $PSNR(F)$ quality.

The proposed $PSNR(F)$, $\widetilde{PSNR}(I)$ and $Q_{F,G}$ metrics aim to amend a limitation of the common $PSNR$ quality assessment by expanding its definition to regions of interest. Through these modifications we show that our constructions account for pixel changes in feature locations and convey quantities which yield a more comprehensive assessment of the degraded image and describe relative degradation in the quality of the observable scene. Hence these modification can be leveraged to maintain/enhance the quality of 'feature' without sacrificing the overall compression. Also, the distortion ratio $Q_{F,G}$ can serve as a parameter to control the relative distortion of the feature space, whereas a standard $PSNR(I)$ assessment lacks this ability.

All of the algorithms were developed in Python and are publicly available for download (GIT repository: https://github.com/Jesse-Redford/Adpative-SVD.git). All computations were performed on the same machine with the following specifications. Intel Core i7 CPU (2.5Ghz), and 16GB DDR3 memory.

(a) Ground Truth

(b) Uniform block $CF : 3.25$, $PSNR(F) : 36.44$, $PSNR(G) : 36.51$, $\widetilde{PSNR}(I) : 36.51$, $PSNR(I) : 36.51$, $Q_{F,G} : -0.2\%$

(c) Feature Oriented $CF : 7.76$, $PSNR(F) : 36.44$, $PSNR(G) : 26.91$, $\widetilde{PSNR}(I) : 28.14$, $PSNR(I) : 27.44$, $Q_{F,G} : 26.14\%$

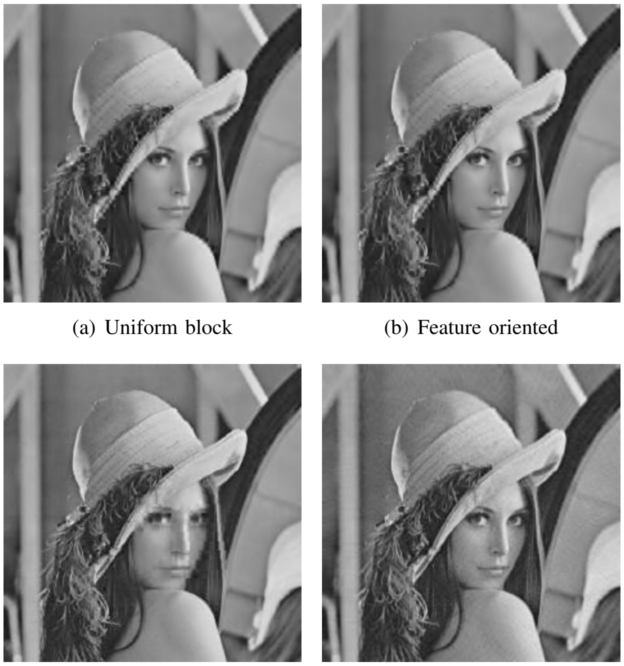
(d) Ground Oriented $CF : 5.12$, $PSNR(F) : 25.39$, $PSNR(G) : 32.33$, $\widetilde{PSNR}(I) : 31.44$, $PSNR(I) : 30.55$, $Q_{F,G} : -27.32\%$

Fig. 2: Assessment of proposed metrics and reconstruction quality of gray scale photo of crowd processed by Algorithm 1. Source image adopted from Bettmann collection via getty images.

V. Performance Comparison

A. Still Image Compression

Figure 3 compares the reconstruction quality for global, uniform block, feature-oriented, and ground-oriented compression on an 8bit 512x512 gray scale image on Lenna, a standard test image widely used in the field of image processing since 1973 [16].



(a) Uniform block (b) Feature oriented

(c) Ground oriented (d) Global SVD

Fig. 3: Compressed images of Lenna.: (a) uniform block; (b) feature oriented; (c) ground oriented; and, (d) global SVD compression.

Table 1 provides performance measures for various compression approaches given in Figure 3.

TABLE I: Compression Performance Still Image of Lenna

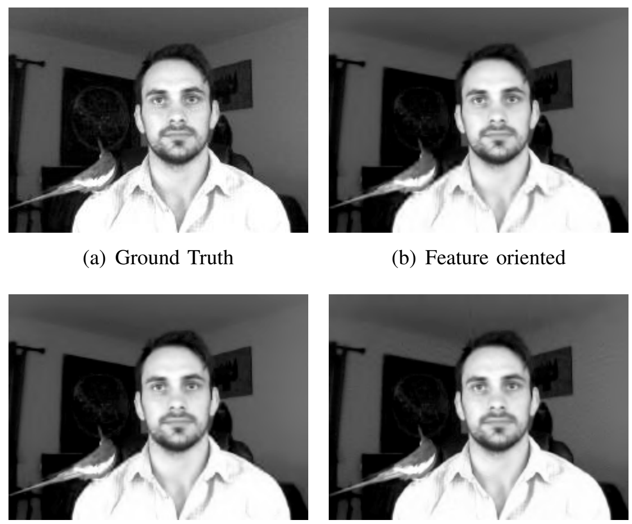
Classification	CF	$Q_{F,G}$	$\widetilde{PSNR}(I)$	$PSNR(F)$
(a) Uniform Block	3.88	1.2 %	30.3	30.97
(b) Feature-Oriented	3.7	6.21 %	30.65	34.59
(c) Ground-Oriented	2.7	-15.27 %	33.05	25.91
(d) Global	3.65	-0.86 %	32.57	32.07

[a]Measured quantities for Figure 3.

From Figure 3, the global method produces a compressed image with additional artificial grainy texture, especially on the face region. Uniform block is seen to have lower $PSNR$ quality in the face region, where as for the feature-oriented these blocking artifacts appear only in the ground region. For ground-oriented compression, we see similar blocking artifacts in the feature region as expected. From Table 1, the performance of each approach ranked in terms of $PSNR$ follow the order global, ground-oriented, feature-oriented, and uniform block compression. Clearly, the feature-oriented approach provides the optimal feature quality while maintaining the second highest compression factor. The feature-oriented approach has $Q_{F,G}$ value of 6.21%, so the relative distortion is within the suggested tolerance. One could argue that the ranked assessment of each scheme in terms of $Q_{F,G}$ has better compliance to the rank given by a human visual assessment of the images in Figure 3. We can conclude that, the FC-SVD algorithm achieves a CF of 3.7 with a $PSNR(F)$ of 34.59, whereas the uniform block compression achieves a slightly higher CF of 3.88, but with much lower $PSNR(F)$ quality.

B. Applications for Dynamic Video Processing

We evaluate the proposed scheme in a live stream application, where each method of compression is used between an 8bit 480×640 video stream and a H.265 encoder, also known as high-efficiency-video-coding (HEVC) [17]. OpenCVs face-recognition model is used to evaluate each frame of the video stream, defining the ROI's for FC-SVD compression algorithm. For other cases, frames are compressed without leveraging the known locations of ROI's. A block size of 32×32 is used for both uniform and feature-oriented implementations. Figure 4 compares each method of compression for a single image frame of the video.



(a) Ground Truth (b) Feature oriented

(c) Uniform (d) Global

Fig. 4: Comparison of compressed video frame: (a) ground truth, h.265 encoding without SVD compression; (b) h.265 with feature-oriented; (c) h.265 with uniform block rank; and, (d) h.265 with global

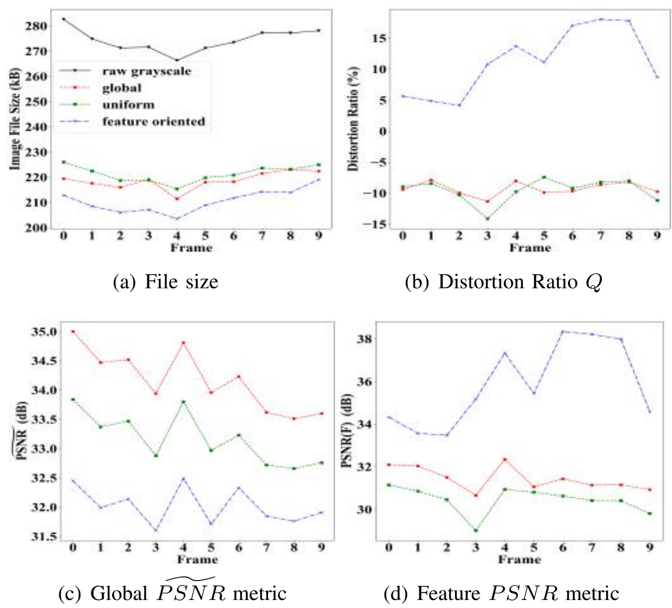
Table 2 provides file sizes of the 2 second video, and average quantities, CF, $\widetilde{PSNR}(I)$, $PSNR(F)$, $Q_{F,G}$ for the series of image frames.

TABLE II: Compression Performance on Live Video Stream

Classification	**Video size in kB**	CF	$\widetilde{PSNR}(I)$	$PSNR(F)$	$Q_{F,G}$
(a) Ground Truth	416	-	-	-	-
(b) Feature-Oriented	93	4.57	32.02	35.85	11.20 %
(c) Uniform Block	99	3.94	33.17	30.45	-9.55 %
(d) Global	101	4.22	34.17	31.44	-9.26 %

[a] Averaged quantities of all 10 frames for Figure 4.

From Table 2, we see that preprocessing each video frame with SVD, regardless of the approach is able to reduce the total video size by approximately 76%. Although, feature-oriented compression results in the smallest file size, in comparison to other methods. In terms of the global image quality $\widetilde{PSNR}(I)$, we see that global SVD outperforms both uniform and feature-oriented SVD. However, we see that the proposed $PSNR(F)$ assessment yields a better objective measure of reconstruction quality in the regions of interest. Figure 5 reports the assessments of each processed frame of the video stream. The subject's face is at a varying distance of approximately 6 to 3 feet from the camera, respectively. Feature-oriented maintains a face space quality ranging between 34 and 38 dB, in addition to having the lowest frame file size for all distances.



(a) File size (b) Distortion Ratio Q

(c) Global $\widetilde{PSNR}$ metric (d) Feature $PSNR$ metric

Fig. 5: Performance Comparison on live video with h.265 encoding: (a) file size in kB of each video frame saved as png; (b) compression factor $Q_{F,G}$ for each video frame; (c) measured $PSNR(I)$ for each video frame , (d) measured feature quality $PSNR(F)$.

The results clearly indicate that feature-oriented compression based on locally adaptive SVD, can keep the same level of compression of global and uniform while maintaining the image quality, especially in the regions of interest. It is also seen that $PSNR(F)$ together with $Q_{F,G}$ can both measure, and quantitatively control the degradation of feature space as well as the amount of artificial distortion.

VI. Comparison to State-of-the-Art

We find that for state-of-the-art results, for example, *"End-to-End Optimized ROI Image Compression"* [14], which is implemented using neural networks. This method and framework achieve much better compression results, however, this framework requires one to train the network with abundant parameters for a particular feature and does not explicitly provide the quality measures for the feature and ground regions, nor the quantity assessments for reducing the distortions. The aim of our method is to contribute a very simple compression framework only based on block SVD and rank-1 truncation that can be adapted easily to support different regions of interests. Moreover, we provide base line examples of image quality defined by the proposed modified $\widetilde{PSNR}$ and relative distortion Q measures. These newly proposed measurements are easy to compute and could provide quantitative assessments for enhancing the compression and evaluating different feature oriented compression algorithms.

VII. Conclusion

A feature compliant compression algorithm FC-SVD is introduced in this paper and a set of objective measures for image compression quality analysis have been presented as well. The suggested FC-SVD algorithm is based on 'feature'-'ground' segmentation and adaptive block SVD compression. Based on a series of experiments, this paper also provided a practical suggestion to adjust the parameters of implementation to improve various quality measures. The proposed algorithm and compression framework is intended to be simple, yet flexible and robust enough to be used in a variety of image processing applications. Our experimental results have shown that this algorithm has the potential to enhance still image as well as video compression over conventional block-based, and global SVD transformations. We will mathematically study these quantities and identify the optimal settings of parameters in the future. We will also extend this framework to machine learning by enhancing the access to processed training data.

References

[1] Nnamdi Ozah and Antonina Kolokolova. Compression improves image classification accuracy. In *Canadian Conference on Artificial Intelligence*, pages 525–530, 2019.

[2] A.J. Hussain, Ali Al-Fayadh, and Naeem Radi. Image compression techniques: A survey in lossless and lossy algorithms. *Neurocomputing*, 300:44–69, 2018.

[3] Xiaolong Ma, Xudong Xie, Kin-Man Lam, Jianming Hu, and Yisheng Zhong. Saliency detection based on singular value decomposition. *Journal of Visual Communication and Image Representation*, 32(C):95–106, 2015.

[4] M.Z Aziz and B Mertsching. Fast and robust generation of feature maps for region-based visual attention. *IEEE Transactions on Image Processing*, 17(5):633–644, 2008.

[5] Abhiram Ranade, Srikanth S Mahabalarao, and Satyen Kale. A variation on svd based image compression. *Image and Vision computing*, 25(6):771–777, 2007.

[6] Nidhal K El Abbadi, Adil Al Rammahi, Dheiaa Shakir Redha, and Mohamed AbdulHameed. Image compression based on svd and mpq-btc. *JCS*, 10(10):2095–2104, 2014.

[7] Francesca De Simone, Daniele Ticca, Frederic Dufaux, Michael Ansorge, and Touradj Ebrahimi. A comparative study of color image compression standards using perceptually driven quality metrics. In *Applications of Digital Image Processing XXXI*, volume 7073, page 70730Z, 2008.

[8] Umme Sara, Morium Akter, and Mohammad Shorif Uddin. Image quality assessment through fsim, ssim, mse and psnr—a comparative study. *Journal of Computer and Communications*, 7(3):8–18, 2019.

[9] Stephen T Welstead. *Fractal and wavelet image compression techniques*, volume 40. Spie Press, 1999.

[10] Mauro Barni. *Document and Image compression*. CRC press, 2006.

[11] Nikolaos Thomos, Nikolaos V Boulgouris, and Michael G Strintzis. Optimized transmission of jpeg2000 streams over wireless channels. *IEEE Transactions on image processing*, 15(1):54–67, 2005.

[12] Xiangjun Li and Jianfei Cai. Robust transmission of jpeg2000 encoded images over packet loss channels. In *2007 IEEE International Conference on Multimedia and Expo*, pages 947–950, 2007.

[13] Renuka G Deshpande, Lata L Ragha, and Satyendra Kumar Sharma. Video quality assessment through psnr estimation for different compression standards. *Indonesian Journal of Electrical Engineering and Computer Science*, 11(3):918–924, 2018.
[14] C. Cai, L. Chen, X. Zhang, and Z. Gao. End-to-end optimized roi image compression. *IEEE Transactions on Image Processing*, 29:3442–3457, 2020.
[15] Vinayak K Bairagi and Ashok M Sapkal. Roi-based dicom image compression for telemedicine. *Sadhana*, 38(1):123–131, 2013.
[16] Susmita Sahu, Himadri Sarma, and Dibya Jyoti Bora. Image segmentation and its different techniques: An in-depth analysis. In *2018 International Conference on Research in Intelligent and Computing in Engineering (RICE)*, pages 1–7, 2018.
[17] Grzegorz Pastuszak and Andrzej Abramowski. Algorithm and architecture design of the h. 265/hevc intra encoder. *IEEE Transactions on circuits and systems for video technology*, 26(1):210–222, 2015.