Locally Adaptive SVD Framework for Feature Oriented Compression

Jesse Redford and *Dr. Xingjie Helen Li*

Abstract

This paper offers a spatially adaptive singular value decomposition (adaptive SVD) framework for flexible reduction of data redundancy commonly found in color images, videos, and their respective grayscale representations. Using a combination of image segmentation and face detection techniques as a pre-processing step, we draw upon developed visual system methodologies, such as visual spatial attention, to derive a compression criterion for controlling the local mode of SVD truncation. We demonstrate the utility of a feature compliant compression algorithm (fc-SVD) and propose a revised peak signal-to-noise ratio (PSNR) quality metric as an extension for its use in various imaging applications. The goal of this investigation is to provide new contributions to image and video processing research involving the analysis of SVD properties.

I. Introduction

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. This form of compression exploits redundancies commonly found in image data in order to increase the storage capacity of devices and optimize the use of network bandwidth. These redundancies typically consist of either Coding redundancy, Inter-pixel redundancy, which results from correlations between pixels, and Physcho-visual redundancies, data that is ignored by the human visual [1]. A well-designed lossy compression scheme can significantly reduce the size of an image before any apparent degradation is noticed by the end user. However, in cases such real-time communication and data transmission, lossy compression is a balancing act between preserving image quality and achieving a desired level of performance. Although there is variety of general image and video compression algorithms available, such as JPEG, JPEG2000, or MPEG standards, not all provide the adaptability that may be required for certain use cases. Helen: rephrase this sentence, for example the JPEG type algorithm are not designed for objective oriented. In this study, we analyze the common SVD image compression schemes and offer suggestions on how adaptive block-based implementations can be leveraged to enhance image quality. Using a combination of image segmentation and face detection techniques as a pre-processing step, we draw upon developed visual system methodologies to derive an objective oriented compression criterion for controlling the local mode of SVD truncation. We formalize our technique as a feature compliant compression algorithm (fc-SVD) and propose a revised PSNR quality metric as an extension for its use in various imaging applications. Our research aims to extent the utility of SVD to Facial recognition, object detection, image segmentation and other processes in which an apparent distinction between available data and information exists [2] [3] [4] [5].

II. Singular Value Decomposition

Singular value decomposition is a robust orthogonal matrix decomposition method which can be preformed on any arbitrary shape of real rectangular or square matrix[1][6]. Suppose we let an image be viewed as matrix $I_{m\times n}$ with $n \leq m$, each pixel being an element of the that matrix, ranging between 0 – $2n$ bits. The singular value decomposition admits the following factorization.

$$I = U\Sigma V^T \tag{1}$$

$$U\Sigma V^T \to \begin{bmatrix} u_1 & u_2 \cdots & u_n \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots\ 0 & \cdots & \sigma_n \end{bmatrix} \begin{bmatrix} v_1^T \\ v_1^T \\ \vdots \\ v_1^T \end{bmatrix}$$

where U is a $m\times m$ unitary matrix, Σ is a $m\times n$ rectangular diagonal matrix with non-negative real numbers along the diagonal, and V is an $m\times n$ unitary matrix. If the imaging matrix $I_{m\times n}$ is real, U and V^T,will be real orthonrmal matrices, representing the left and right singular vectors of the decomposition. The rank r of the matrix I is described by the number singular values, non-zero entries, which appear along the diagonal of the matrix Σ, where $r \leq n \leq m$.

Singular values follow the order

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0 \tag{2}$$

such that the most dominate singular values appear in the first few terms of the diagonal. It is often of interest to compute a low rank-approximation of the matrix I by truncating the singular values after k terms, as shown in equation 3.

$$(\sigma_{k+1} \cdots \sigma_n) = 0$$

This paradigm works well when you can find a representation of the data in which most of the important information is concentrated in just a few components of the decomposition.

III. Image Compression

The use of SVD in image compression has been widely studied. [1] The concept of SVD as a compression tool is described by Equations (3)(4)(5). Where the amount of memory required by the image I is defined by the number of matrix entries used to represent it. For an image with $m \times n$ pixels, representing this image as matrix $I_{m\times n}$, requires $m \times n$ memory locations. As seen in Equation (5), Leveraging the SVD decomposition to approximate the matrix I, yields an alternative representation requiring only k(m+n+1) memory locations to store the image. Another advantage of using the SVD for compression is the property of energy compactions, and its ability to adapt to the local statisical variations of an image. [6]

$$I_{mxn} \approx\ I_k\ =\ U_{mxk}\Sigma_{kxk}V_{kxn}^T = \sum_{i=1}^{k}\sigma_i u_i v_i^T \tag{3}$$

$$m \times n = (m \times k)\,(k)\,(k \times n) = k\,(1+m+n) \tag{4}$$

We define compression as a ratio between the uncompressed representation of the image data to its compressed representation.

$$\frac{mn}{k(m+n+1)} : \frac{mn}{mn} \tag{5}$$

SVD is considered a lossy type of compression due to condition given by Equation (6), which describes the minimum rank approximation-k required to achieve compression.

$$CF(k) = \frac{m \times n}{k\,(1+m+n)} \quad \geq 1 \tag{6}$$

It should be noted that (3)(4)(5)(6) outline a theoretical compression criteria for SVD, in practice the effective compression ratio is evaluated as an image's uncompressed file size divided by the compressed file size in bytes. For our results we report Equation 5 as CF, and the effective compression as percent space savings denoted as SS, in Equation().

$$SS\ =\ \left(1 - \frac{compressed\ file\ size_{(bytes)}}{uncompressed\ file\ size_{(bytes)}}\right) *\ 100 \tag{7}$$

IV. Image Quality

As described in Section (3), Image compression is minimizing the size in bytes of a graphics file without degrading the quality of the image to an unacceptable level.(Add reff) For SVD compression an emphasis is placed on the determination of the optimal k-rank approximation Ik for some matrix I. This process often consist of minimizing there difference according to a loss function described by Equation (9).

$$I_k = argmin_{I'_k}||I - I'_k|| \tag{8}$$

Selecting a loss function which yields the optimal k-rank approximation may often be domain specific. However, for image compression selecting a loss function which agrees with Human Visual Perception is often advantageous. PSNR or peak-signal-to-noise-ratio is debated to give the best error estimation for measuring the reconstruction quality of a compressed image.[3] PSNR is given by Equation(9).

$$\mathrm{PSNR}\ =\ 10\mathrm{log}_{10}\left(\frac{\mathrm{MAX_I^2}}{\frac{1}{\mathrm{mn}}\sum_{\mathrm{i=0}}^{\mathrm{m-1}}\sum_{\mathrm{j=0}}^{\mathrm{n-1}}\left[\mathrm{I\,(i,j)} - \mathrm{I_k\,(i,j)}\right]^2}\right) \tag{9}$$

Where m and n is the number of rows and columns of the image matrix and $\mathrm{MAX_I}$ is determined using Equation (10). Where n represents the number of bits used to represent the image.

$$\mathrm{MAX}_I = 2^n - 1 \tag{10}$$

Typical PSNR values for lossy image compression range between 30 and 50db for 8 bit images, 60 to 80db for 16 bit images. However, wireless transmission quality often accepts PSNR values ranging between 20 and 25db. [3] Another, error metric often used to evaluate image quality is given by Equation (11), denoted as RMSE for root-mean-square-error.

$$RMSE = 1mni = 0m - 1j = 0n - 1[I(i,j) - Ik(i,j)]2 \tag{11}$$

Note PSNR shares a positive correlation with image quality as opposed to RMSE, which is shares an inverse correlation, meaning that a low RMSE value indicates higher quality image.

V. Block Extensions For SVD Compression

Common SVD compression schemes often entail block based implementations, which often yields better compression performance and reduces processing time. Golub-Van loan gave Equation (12) as the computation complexity for computing the SVD,[add reff] denoted as TC for time complexity in our results, to evaluate the computation cost of the decomposition.

$$TC\ =\ O(m^2n)\ ,\ m\ >\ n \tag{12}$$

We extend Equation (5) and the performance criteria outlined in section2 for cases in which a series of sub-matrix decompositions are used to represent the image matrix. Given the image matrix I with dimensions M by N, and an area which is divisible by a smaller matrix of dimension m, n, the number of sub-matrices required to represent the matrix can be expressed as

$$\frac{M}{m_i} * \frac{N}{n_i} \in \mathbb{Z} \tag{13}$$

Where

$\frac{M}{m_i}$ is the required number of row iterations

$\frac{N}{n_i}$ is the number of column iterations.

Evaluation of a block based approach is assessed using Equations (14) and Equation (15).

$$TC = \left[\max\left(m_{b_i}, n_{b_i}\right)^2 \min\left(m_{b_i}, n_{b_i}\right)\right] \frac{M}{m_i} * \frac{N}{n_i} \qquad (14)$$

$$CF = \frac{MN}{\left[k_i\left(1 + m_{b_i} + n_{b_i}\right) * \left(\frac{M_i}{m_{b_i}} \frac{N_{s_i}}{n_{b_i}}\right)\right]} \qquad (15)$$

We study the case of k=1 for all divisible factors of the image matrix. Thus, the reconstruction quality, compression, and computation cost is governed by the block size dimensions and the relative pixel complexity residing within each of our sub-matrix approximations. Figure 1 represents examples of image quality for a rank-1 update for using different divisible factors to compress the GS image of ().

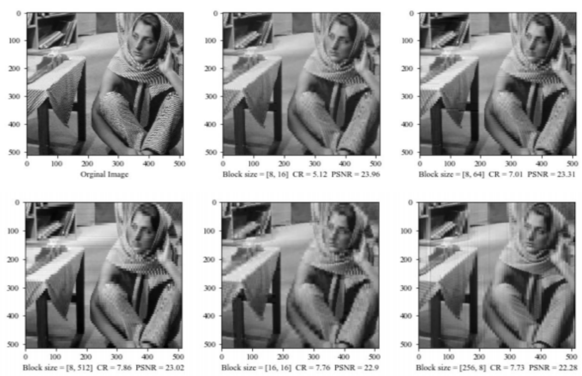

Figure 1: Compressed images using various rank-1 submatrix approximations. Blocksize, compression factor, and PSNR

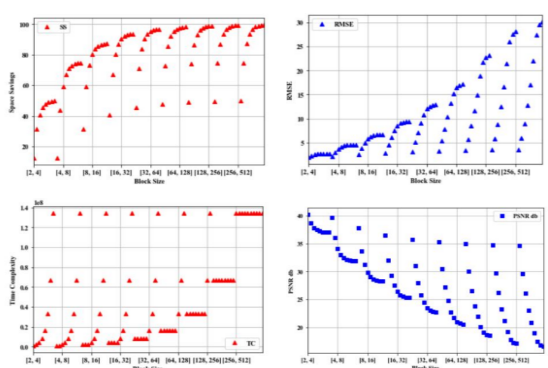

Figure 2: Comparison compression performance vs blocksize for rank1-update (a) Space Savings versus block-size, (b) root-mean-square-error versus block-size, (c) time complexity versus blocksize, (d) peak-signal-to-noise-ratio versus block-size.

Figure 2 reports the recorded compression performance using all divisible factors to compress a 512x512 image of Lena. Figure 2 clearly displays the viability and effectiveness of various block-sizes for different quality and compression criteria.

VI. SVD Subspace Approximations

Given an asymmetrical case described by Figure 2, where an image is partitioned into multiple sections, introduce Equations (16)(17) as extensions for the performance criteria outlined by (6)(12)(14)(15).

$$TC = \sum_{s=1}^{S_n} \left[\max\left(m_{b_i}, n_{b_i}\right)^2 \min\left(m_{b_i}, n_{b_i}\right) \left(\frac{m_{s_i}}{m_{b_i}} \frac{n_{s_i}}{n_{b_i}}\right)\right] \qquad (16)$$

$$CF = \frac{MN}{\sum_{s=1}^{S_n} \left[k_i\left(1 + m_{b_i} + m_{b_i}\right) * \left(\frac{m_{s_i}}{m_{b_i}} \frac{n_{s_i}}{n_{b_i}}\right)\right]} \qquad (17)$$

Where M,N are the dimensions of the image matrix, and S is a unique partition with dimensions ns ms divisible by a smaller matrix with dimensions mb nb. These extensions desribe partion cases such as the one given in Figure().

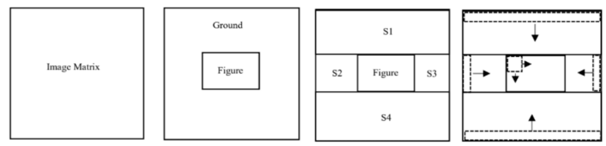

Figure 3: Visual representation of figure-ground segmentation of image matrix

In order to better accommodate geometric deviations which appear in an image figure, we suggest the use of a convex hull algorithm to refine boundary conditions which enclose ROIs.

VII. Leveraging ROIs and HSV Characteristics

In this section we outline our contributions for developing a feature compliant compression scheme inspired by the human visual system HVS. Our results suggest 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. We attempt to leverage this concept by defining a ROI region-of-interest(s) within image frames, which we refer to as the image's "Figure", referencing the remaining portions of the image as "Ground". This distinction tailors compression rates in the spatial domain in order to enhance compliance with the HVS by prioritizing computational attention to different regions of an image or video stream.

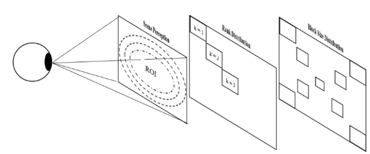

Figure 4: Example of our proposed compression scheme for enhancing compliance with HVS human visual system

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 $F(xi, yi)$ are available to construct boundary conditions.

Algorithm 1: Feature Compliant Compression

load: detection model Θ
set: error type ξ compression factor Φ tolerance τ
initialize: $\{\chi_i\}_{i=1}^{\infty} \leftarrow image\ stream$
for $i = 1$ **to** ∞ **do**
  $\chi = \chi_i \leftarrow read\ imageframe$
  $F(x_i, y_i) = f(\Theta, \chi) \leftarrow detect\ features\ points$
  $S[\mathrm{m}_i, n_i] = (M|m', N|n') \leftarrow evaluate\ divisibility$
  for m_i *and* n_i *in* χ **do**
    $\chi_b = \chi[x : x + m, y : y + n]$
    if χ_b not in $F(x_{min} : x_{max}, y_{min} : y_{max})$
    $\Psi = 1$
    otherwise
    $\Psi =$low rank
    $\chi_b' \leftarrow$svd(χ_b,k=Ψ)
  check ξ,Φ,τ = True $\leftarrow evaluate\ approximation$
  $\chi_{out} = \chi_b' \leftarrow transmit\ compressed\ image$
  otherwise $S = S + 1 \leftarrow update\ block$ m_i,n_i

We propose Equation(18) and Equation (19) as extensions for evaluating feature compliant compression, for cases which the regulation of error in both figure and ground portions of the image are of intrest. This process consist of computing the relative quality in both figure and ground portions of the image.

$$PSNR_T = \frac{1}{\sum_{i\ =\ 1}^{S\ =\ F_i + G_i} \quad w_s \frac{1}{PSNR(s_i)}}\ , w_s = \ \frac{MN}{M\,(s_i)\,N(s_i)} \tag{18}$$

$$PSNR_G = \sum_{i}^{S_i} \frac{M\,(s_i)\,N(s_i)}{MN - \ M\,(F)\,N(F)} \tag{19}$$

VIII. Experimental Results

Comparisons of the reconstruction quality of gray scale image on Lenna. standard homogeneous block compression scheme vs adaptive

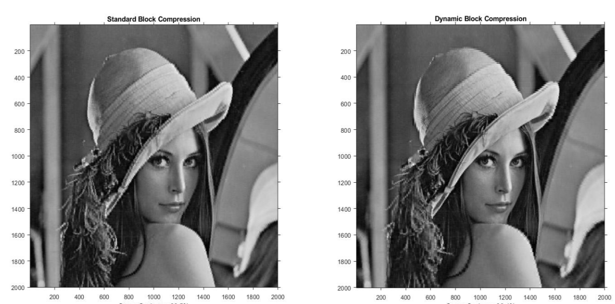

Figure 5: Left standard block based implantation. Right adaptive block

In the following we evaluate the proposed fcSVD algorithm and compare the performance to global SVD and uniform rank1-block SVD. All of the algorithms were developed in Python and are publicly available for download (GIT repository: https). All computations were performed on the same machine with the following specifications. Intel Core i7 CPU (2.5Ghz), and 16GB DDR3 memory. Localized SVD compression is used as a preprocessing step between a raw 8bit camera stream and a H.265 encoder.H.265. is also known as (HEVC) or high-efficiency-video-coding. The raw grayscale feed is passed to a global, block, and adaptive block compression scheme. Both the global, and block schemes compress the incoming gs frame. Each compressed frame is evaluated using OpenCV face recognition as the means of detection, with the exception of the adaptive block scheme which utilizes face detection before compression. Coordinates of face locations are leveraged to adaptively regulate compression rates throughout the image.

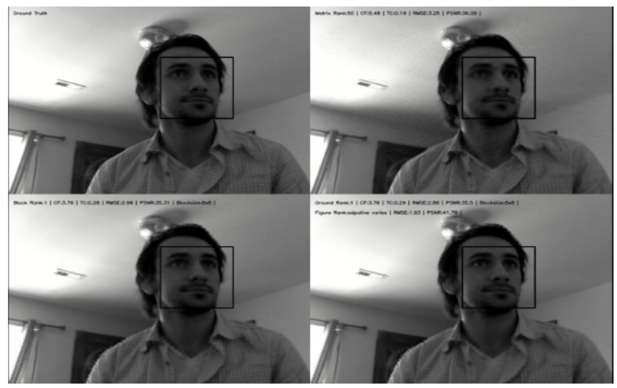

Figure 6: Figure1: Top left: ground truth, h.265 encoding without SVD compression. Top right: h.265 with global-SVD Bottom left: h.265 with rank 1 - block SVD Bottom right:h.265 with adaptive block SVD

IX. Discussion

The suggested figure-ground segmentation and compression methods evaluated in this paper are intended to be simple,yet flexible and accurate enough to be used in a variety of image processing applications.The proposed compression method for rank approximation differs from most low-rank method. As the suggested method uses only 1 eigenvalue for each submatrix approximation. Thus compression and quality rates are governed by the dimensions of our sub-block matrices.The proposed figure-ground segmentation methods also allow for compression via dimension reduction by simply extracting the image figure I_F' with dimensions M_F', N_F'. Given that the area enclosed by our figure matrix I_F is less than that of the input matrix I. Our Experimental results shown that Adaptive SVD has the potential to be enhance video compression when used as a pre-processing step for h.264 or h.265 encoding. Future work seeks to further evaluate the potential of an feature compliant compression scheme, and futher extensions which can be leveraged in surveillance applications which have limited on board-storage and bandwidth.

X. Acknowledgments

Special thanks to Dr. Xingjie (Helen) Li for her guidance and support throughout this project.

References

[1] 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.

[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] HS Prasantha, HL Shashidhara, and KN Balasubramanya Murthy. Image compression using svd. In *International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007)*, volume 3, pages 143–145. IEEE, 2007.