Adaptive Image Compression Using Saliency and KAZE Features

Siddharth Srivastava, Prerana Mukherjee and Brejesh Lall
Department of Electrical Engineering, Indian Institute of Technology, Delhi, India
{eez127506, eez138300, brejesh}@ee.iitd.ac.in

***Abstract*—In this paper we present a novel framework for compressing images using saliency maps and KAZE features. The method involves adapting the quality factor in JPEG compression scheme for each block instead of using the same quality factor for the entire image. This is achieved by adapting JPEG quality parameter based on visual saliency and KAZE keypoints. Subsequently, a piecewise function is used to compress the least important image blocks with higher compression ratio while maintaining the overall perceptual quality and avoiding blocking artifacts. This work introduces use of KAZE keypoints for image compression for the first time in literature. We show that the proposed method outperforms the JPEG compression using PSNR and FSIMc evaluation measures especially at high compression ratios.**

Keywords- JPEG Compression, Saliency, KAZE, Quality Parameter

I. Introduction

This paper describes an image compression scheme based on adaptively selecting i.e. boosting the quality parameter for each block during JPEG image compression. The boosting is achieved by utilizing the saliency map of [1] and KAZE features [2]. Saliency helps in allocating more number of bits to the regions with high perceptual content as compared to those with lower or insignificant perceptual content. KAZE keypoints, help to identify the regions of sharp changes in the images by providing keypoints along the boundaries in the image. This approach shows a lot of potential as the results indicate significant gains over the standard JPEG. This is achieved with minimal change in the JPEG standard bit stream making the approach backward compatible. We evaluate our proposed scheme using Peak Signal-to-Noise Ratio (PSNR) and Feature Similarity Index (FSIMc) [3] evaluation measures.

In recent times, saliency driven compression methods have gained popularity. Authors in [4] utilize a depth preserving saliency map based compression scheme. They calculate the blur amount from the Difference of Gaussian images, which are used as a cue for indicating the depth in the original image. They further adapt the JPEG algorithm to use a block based quality parameter derived from the saliency of each block relative to the entire image. In [5] the authors form a framework for saliency computation using phase spectrum of quaternion Fourier transform (PQFT) of the image. It defines the regions with less homogeneity and less periodicity as salient regions. They use the spatio-temporal PQFT saliency map to form a hierarchical framework for the image and obtain an efficient coding strategy for images and videos. In [6], authors use the eye tracking information and integrates them with the distortion metrics in the JPEG compression. Saliency based methods have also found application in video compression. The authors in [7] propose to perform non-uniform anti-aliasing. The non-linearly downscaled image can be compressed using any compression scheme. In [8], the authors use the saliency model of [9] to find region of interests with high attentional selection in order to reduce the compression artifacts in corresponding regions.

The previous discussion illustrates that saliency based compression techniques yield better results when augmented with stronger selection criteria for salient blocks. To this end, we introduce KAZE keypoint based block selection criteria to strengthen the block based quality parameter in the JPEG compression scheme. Previous works using local feature extraction techniques for compression [10][11] involve very high computational resources along with use of machine learning algorithms. They construct a low resolution image from the original image along with extracting Scale Invariant Feature Transform (SIFT) [12] features. The low resolution image and SIFT features are transmitted to a remote location (usually a cloud) where the features are used to obtain candidate patches from a large database of patches. The patches corresponding to the features undergo geometric verification and are combined to reconstruct the complete image. These algorithms are not suitable for local storage or on devices with zero or low bandwidth availability. Moreover, such techniques do not generalize well since they are dependent on availability of the patches of the image being compressed on the cloud. Our paper addresses these issues and the contributions can be summarised as follows:

- We specifically enhance the regions with objects by using the saliency map of [1] which characterizes the objects as per human perceptual system. This is motivated by the work in [4]. In our work we use a different saliency map, and introduce local features and a piecewise adaptive strategy for calculation of quality parameter.
- We introduce KAZE keypoints for evaluating the content in a block. This knowledge is used to appropriately adapt the Quality Parameter (QP) resulting in a higher quality reconstructed image (for the same encoded file size). To the best of our knowledge, this is the first work to utilize a keypoint based strategy to achieve compression based on the original image itself.

Rest of the paper is organized as follows. Section II provides an overview of the techniques used while Section III discusses the methodology. The results are discussed in Section IV. Finally, Section V concludes the paper.

978-1-5090-1746-1/16/$31.00

II. Background

In this work, the saliency map is computed using the technique proposed in [1]. We briefly review the approach here. The technique involves preprocessing the input image using non-linear anisotropic diffusion filtering to preserve the edge and smoothen the irrelevant details. The saliency map is composed of a weighted combination of various features obtained from the input image and is given as,

$$saliencyMap = w_1 * Cue_L + w_2 * Cue_G + w_3 * Cue_R \quad (1)$$

where Cue_L represents local cues (color, intensity, orientation, depth, motion), Cue_G represents global cues (global contrast, spatial sparsity) and Cue_R represents rarity (PQFT) features. The weights are calculated as,

$$w_i = \frac{\sigma_i}{\sum_{i=1}^{N} \sigma_i} \quad (2)$$

where w_i, $i = 1, 2, 3$ is the weight of the i^{th} map, N are the total number of maps and σ_i is the variance of the i^{th} map.

Although the authors evaluate the effectiveness of the saliency map for segmentation, the combination of various cues makes it a suitable candidate for mapping objects to their saliency levels. Since, it utilizes the pre-processed saliency map as an input for segmentation, the performance metrics (precision and recall rates) improve. The saliency map for a test image is shown in Figure 1 (b).

The other cue we use is boundary keypoints and their strength. KAZE is a recently proposed feature extraction technique which is based on non-linear scale space. It gives stronger response around the boundaries thus preserving the edge information (Figure 1 (c)). KAZE makes use of the conductivity functions which suitably assigns more smoothening to the regions on both sides of the edges, thus promoting more contrast. The non-linear scale space construction in KAZE is given as follows,

$$\frac{\partial L_s}{\partial t} = div \left\{ \left(c\left(u, v, t\right) . \nabla \left(L_s\right) \right) \right\} \quad (3)$$

where *div* and ∇ are divergence and gradient operators respectively, *c* is the conductivity function, L_s is the Gaussian smoothed original image and *t* is scale parameter.

III. Proposed Methodology

The workflow of our proposed method for compressing an image is shown in Figure 2 (a). The saliency map and KAZE keypoints are calculated from the original image. This is followed by calculation of saliency and keypoint strength parameters for each 8x8 block used during JPEG compression (described in subsections III-B and III-C). The parameters thus calculated are further utilized to derive a suitable quality parameter for each block (subsection III-D). To make the proposed approach backward compatible these QP values are passed to a standard implementation of JPEG. In this paper, the block based quality parameters are stored in form of a look-up table and are used during decompression (Figure 2 (b)). This can be incorporated in the JPEG bit stream for complete interoperability. We discuss and describe the steps outlined above in detail in the following subsections.

A. Saliency Map Computation

As discussed in Section II, saliency map is computed using the method of [1] on the image pre-processed with non-linear anisotropic diffusion filtering. The considered saliency map is suitable to be used for image compression techniques having adaptive block based compression due to following reasons. First, it is able to preserve edge information i.e. it accentuates the edges while blurring out other regions. Second, it incorporates both spatial and frequency cues having uniform saliency across a region. This helps in avoiding blocking artifacts due to different quality parameters for various regions, hence maintaining perceptually uniform compression for each region.

B. Boosting Quality Parameter (QP) with Saliency Response

JPEG uses a universal quality parameter for each block. Recently, the authors in [4] have shown that an adaptive quality parameter helps in preserving the regions of high visual importance over less significant regions, thus resulting in less perceptual degradation after compression in the salient regions. The adaptive quality factor for a block i is calculated as,

$$Q_{Sal}(i) = Q_{JPEG} * salBoost(i) \quad (4)$$

where $Q_{Sal}(i)$ is saliency driven quality factor for a block i, Q_{JPEG} is the universal quality factor used in JPEG, $mean(block\_sal(i))$ is the mean of the saliency values in the corresponding block. $mean(salmap)$ is the mean of the computed saliency map and $salBoost$ is the factor which governs the adaptation of Q_U as per the saliency of the block with respect to the image and is given below:

$$salBoost(i) = \sqrt{\frac{mean(block\_sal(i))}{mean(salmap)}} \quad (5)$$

As can be observed in Equation 4, a major advantage of this scheme is that the adapted quality parameter Q_{Sal} is dependent upon the global quality parameter Q_{JPEG}, providing flexibility for adjustment as per the requisite level of image quality.

C. Boosting Quality Parameter with KAZE Keypoint Response

In this work, we impose an additional quality constraint based on boundary strength. This is achieved using KAZE as described in this subsection. For each block, the quality factor is further adapted to the image content by incorporating the KAZE feature responses using the piecewise adaptive Algorithm 1 (Subsection III-D). KAZE keypoints have stronger response along the boundary of the objects in an image. A key observation is that the strength of the keypoint signifies relative importance of a small neighbourhood around it as compared to other keypoints. This is particularly important in case of high compression ratios where blocking artifacts are frequent. Since the KAZE keypoints are computed over a nonlinear scale space, at low compression ratios the blocks with high keypoint

Fig. 1: a) Original Image b) Saliency Map obtained by [1] c) KAZE keypoints obtained by [2]

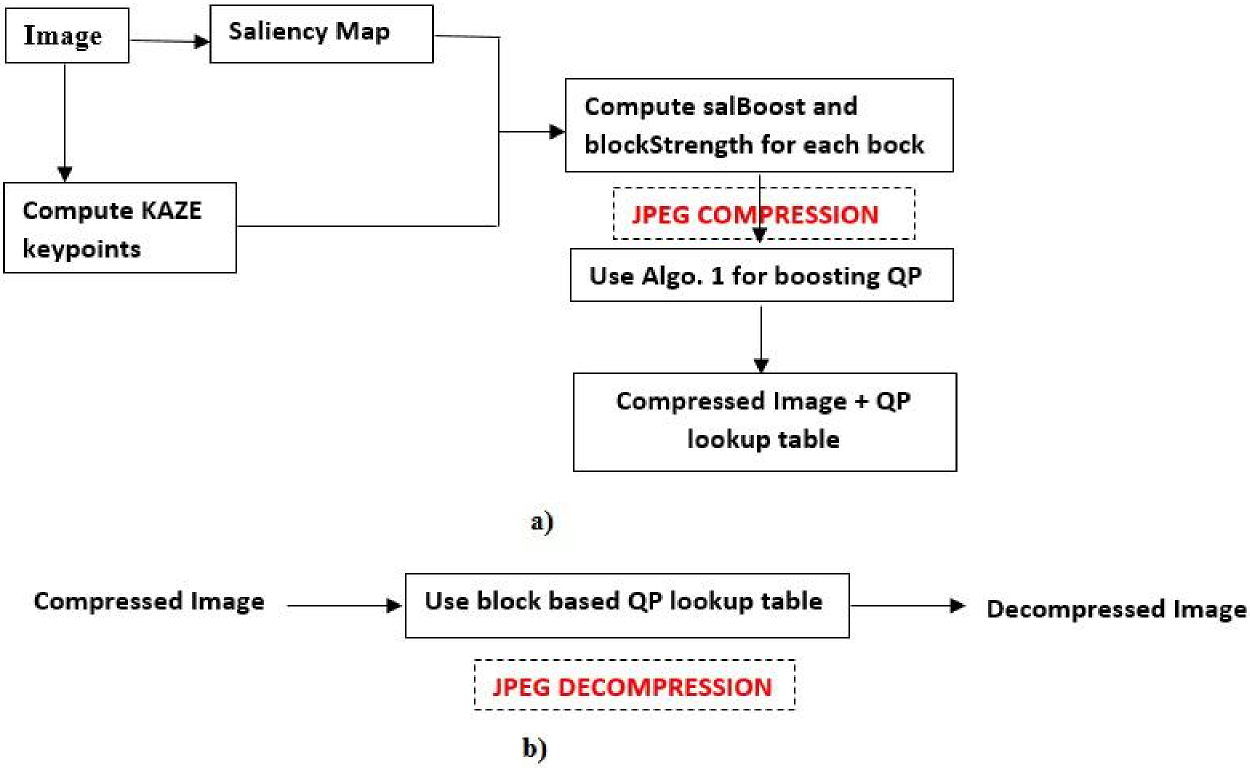

Fig. 2: a) Compression Scheme b) Decompression Scheme

response are assigned more number of bits by further boosting quality parameter over that of saliency based boosting factor described in previous subsection. The strength of a block based on keypoint responses is given by $blockStrength$ which is defined as

$$blockStrength(i) = \sqrt{\frac{mean(block\_response(i))}{mean(image\_response)}} \tag{6}$$

where $mean(block\_response(i))$ is the average of the response of the keypoints in i^{th} block while $mean(image\_response)$ is the average of the keypoint response for the image.

D. *Piecewise Adaptive Quality Parameter*

The final quality parameter ($boostedQP$) is calculated based on a piecewise adaptation of quality parameter for each block. The piecewise nature serves two purpose. First, it allows for uniformity among the quality parameters assigned to blocks lying between boundaries of objects or regions within objects. Hence allowing for smooth transition in the region between blocks. Second, it compresses the perceptually insignificant blocks (less salient and with no or less keypoints) more as compared to perceptually significant blocks (highly salient or with strong response keypoints). This results in increasing the perceptual quality at the same compression ratio. The method is described in Algorithm 1. The values of $\alpha, \beta 1$ and $\beta 2$ govern the amount of compression for each block. They can be chosen interactively for best performance, but in general a fixed value provides consistent results. The piecewise function enhances a block $\beta 2$ times the saliency based quality parameter (Q_{sal}) if the mean response of KAZE keypoints in that block is significantly higher than that of all the keypoints in the image. This condition ensures that blocks with high saliency and having edges maintain there quality. Additionally, if block does not have any KAZE keypoints and the relative saliency of the block is lesser than the image, then the block under consideration is compressed more ($\beta 1$ times). This indicates that the block neither contains any high frequency components nor any visually salient information. The parameter α is used to define a range for mean response of KAZE keypoints within a block and accordingly adjusts the boosted quality parameter ($boostedQP$).

Algorithm 1 Algorithm for Piecewise Adaptive QP

```
procedure BOOSTING–QP^a
    Q_Sal = Q_JPEG * salBoost
    M_R = mean(image_response) [ref. Sec III-C]
    B_S = blockStrength [ref. Equation 6]
    if (B_S == 0 and salBoost < 1)
        boostedQP = β1 * Q_Sal;
    else if (B_S == 0 and salBoost ≥ 1)
        boostedQP = Q_Sal;
    else if (B_S < α * M_R)
        boostedQP = (1 − α) * Q_Sal;
    else if (B_S ≥ α * M_R and B_S < (1 − α) * M_R)
        boostedQP = Q_Sal;
    else if(B_S ≤ (1 − α) * M_R and B_S ≤ (1 + α) * M_R)
        boostedQP = (1 + α) * Q_Sal;
    else
        boostedQP = β2 * Q_Sal;
    end if
end procedure
```

[a] $0 < \alpha \leq 0.5$, $\beta 1 < 1$ and $\beta 2 > 1$

IV. Results

The results are obtained using the Miscellaneous category of database from USC-SIPI database [13]. This category consists of a total of 44 images including both color and grayscale images with varying resolutions. The evaluation is performed by keeping the compression ratio same for all the images in the database and calculating the mean PSNR and FSIMc for the dataset. The results for PSNR and FSIMc are shown in Figure 3 and 4 respectively. The size of the compressed image involves both the size of the image and the size of the quality factor look-up table.

As can be observed, the proposed method outperforms the JPEG compression for both PSNR and FSIMc values. The PSNR is nearly 2-4 dB higher than the standard JPEG

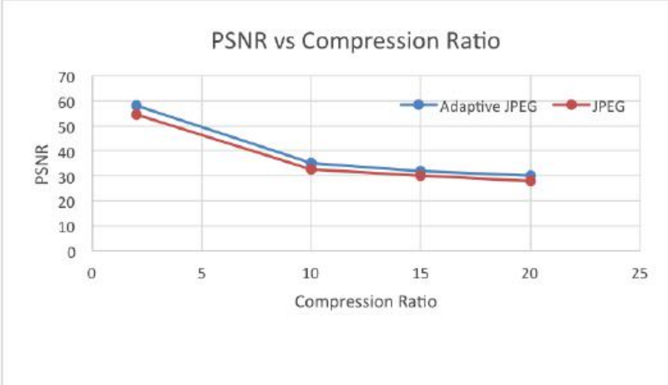

Fig. 3: PSNR vs Compression Ratio

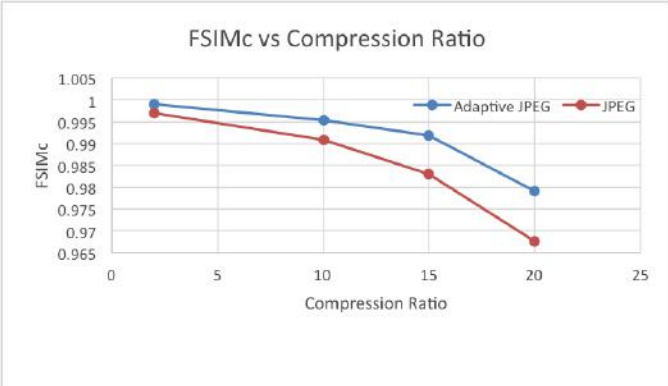

Fig. 4: FSIMc vs Compression Ratio

compression scheme. On the other hand, the FSIMc shows that our method clearly outperforms the JPEG when the perceptual quality of the compressed image is considered. A key observation is that at high compression ratio there is relatively slower drop in quality as compared to JPEG. This demonstrates the effectiveness of our method to retain significantly better perceptual quality even at high compression rates. An example of the images compressed using JPEG and our scheme is shown in Figure 5. The images are compressed at a high compression ratio in order to bring our the differences in terms of the perceptual quality. As can be seen JPEG compression introduces blocking artifacts at such compression ratios while our scheme has negligible artifacts.

V. Conclusion

This paper proposed a JPEG compatible novel visual saliency and KAZE keypoint based compression scheme. The paper also introduced boosting of block based quality factor using KAZE keypoint responses for the first time in literature. The paper also discusses a piecewise function to adapt the quality factor to maintain good perceptual quality at high compression ratios. The effectiveness of the proposed scheme has been analyzed with PSNR and FSIMc evaluation measures and the proposed scheme outperformes the standard

Fig. 5: Comparison at a compression ratio of 15 a) Original Image b) Compressed Image using JPEG Compression Scheme c) Compressed Image using Our Proposed Compression Scheme

JPEG compression while retaining the potential for ensuring interoperability.

REFERENCES

[1] P. Mukherjee, B. Lall, and A. Shah, "Saliency map based improved segmentation," in *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 1290–1294.

[2] P. F. Alcantarilla, A. Bartoli, and A. J. Davison, "Kaze features," in *Computer Vision–ECCV 2012*. Springer, 2012, pp. 214–227.

[3] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "Fsim: a feature similarity index for image quality assessment," *Image Processing, IEEE Transactions on*, vol. 20, no. 8, pp. 2378–2386, 2011.

[4] M. T. Khanna, K. Rai, S. Chaudhury, and B. Lall, "Perceptual depth preserving saliency based image compression," in *Proceedings of the 2nd International Conference on Perception and Machine Intelligence*. ACM, 2015, pp. 218–223.

[5] C. Guo and L. Zhang, "A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression," *Image Processing, IEEE Transactions on*, vol. 19, no. 1, pp. 185–198, 2010.

[6] J. Redi, H. Liu, P. Gastaldo, R. Zunino, and I. Heynderickx, "How to apply spatial saliency into objective metrics for jpeg compressed images?" in *Image Processing (ICIP), 2009 16th IEEE International Conference on*. IEEE, 2009, pp. 961–964.

[7] F. Zund, Y. Pritch, A. Sorkine-Hornung, S. Mangold, and T. Gross, "Content-aware compression using saliency-driven image retargeting," in *Image Processing (ICIP), 2013 20th IEEE International Conference on*. IEEE, 2013.

[8] H. Hadizadeh and I. V. Bajic, "Saliency-aware video compression," *Image Processing, IEEE Transactions on*, vol. 23, no. 1, pp. 19–33, 2014.

[9] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 11, pp. 1254–1259, 1998.

[10] H. Yue, X. Sun, F. Wu, and J. Yang, "Sift-based image compression," in *Multimedia and Expo (ICME), 2012 IEEE International Conference on*. IEEE, 2012, pp. 473–478.

[11] H. Yue, X. Sun, J. Yang, and F. Wu, "Cloud-based image coding for mobile devices-toward thousands to one compression," *Multimedia, IEEE Transactions on*, vol. 15, no. 4, pp. 845–857, 2013.

[12] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[13] "USC-SIPI Image Database," : http://sipi.use.edu/database/, [Online; accessed 10-Oct-2015].