UNIVERSITÀ DEGLI STUDI FIRENZE

Scuola di Scienze Matematiche, Fisiche e Naturali
Corso di Laurea in Informatica

Tesi di Laurea

TITOLO ITALIANO

TITOLO INGLESE

LORENZO PALLONI

Relatore: *Relatore*
Correlatore: *Correlatore*

Anno Accademico 2014-2015

INDICE

ELENCO DELLE FIGURE

"Inserire citazione"
— Inserire autore citazione

1

INTRODUCTION

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

2

QUALITY METRICS

This chapter investigates various Image Quality Assessment (IQA) metrics commonly used in the current literature to evaluate results and improve performance of super-resolution and compression artifact removal operations. For our purpose, a metric is as good as how much it agrees with the average human judgement. The agreement is often measured with Pearson, Kendall, and Spearman correlation coefficients, while the average human judgement is estimated using sets of images structured for Mean Opinion Score (MOS), Two Alternative Force Choices (2AFC), or Just Noticeable Difference (JND) approaches.

It is difficult to define an objective quantification of results from super-resolution and compression artifact removal operations that reflects human judgement.

Plenty of quality metrics have been proposed in the literature.

Traditional metrics, such as Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR) 2, Structural Similarity (SSIM) [16], and Multi-Scale SSIM (MS-SSIM) [17] have begun to be considered inadequate for Image Quality Assessment (IQA) in recent years.

A promising line of work is based on deep-learning features extracted from well-trained CNN models like AlexNet and VGG-16. FID [3], LPIPS [19], LPIPS-Comp [11], E-LPIPS [5], and DISTS [2] are all examples of deep-learning-based metrics.

Another approach that enphasizes the importance of edges in restored images is ERQA, introduced in [6], and improved in [7], a metric based on the Canny edge detector. - SFSN (Structural Fidelity versus Statistical naturalness [21] So far, only Full-Reference IQA (FR-IQA) measures have been mentioned, but No-Reference IQA (NR-IQA) metrics - that evaluate processed images without their originals - are commonly used, and some examples are BRISQUE, NIQE, PIQE, and CONTRIQUE.

LIVE, TID2008, CSIQ, TID2013 are examples of FR-IQA datasets, while AVA, LIVE In the Wild are NR-IQA datasets, that is they assess the quality of an image by itself, without a reference image.

Let y be an image and $\hat{y}$ be a distorted version of y. A measure of similarity between the two given images is the Peak Signal-to-Noise Ratio:

$$\mathrm{PSNR}(y, \hat{y}) := 10 \cdot \log_{10} \left\{ \frac{\mathrm{MAX}(y)^2}{\mathrm{MSE}(y, \hat{y})} \right\}, \tag{2.1}$$

where $MAX(y)$ is the maximum value of y, and $MSE(y, \hat{y}) := \|y - \hat{y}\|_2^2$.

2.1 SSIM

Let $y = \{y_i | i = 1, 2, \ldots, N\}$ and $\hat{y} = \{\hat{y}_i | i = 1, 2, \ldots, N\}$ be two discrete non-negative signals that have been aligned with each other, and let μ_y, σ_y and $\sigma_{y\hat{y}}$ be the mean of y, the variance of y, and the covariance of y and $\hat{y}$, respectively. Approximately, μ_y and σ_y can be viewed as estimates of the luminance and constrast of x, and $\sigma_{y\hat{y}}$ measures the tendency of y and $\hat{y}$ to vary together, thus an indication of structural similarity. In [16], the luminance, constrast and structure comparison measures were given as follows:

$$l(y, \hat{y}) = \frac{2\mu_y \mu_{\hat{y}} + C_1}{\mu_y^2 + \mu_{\hat{y}}^2 + C_1} \tag{2.2}$$

$$c(y, \hat{y}) = \frac{2\sigma_y \sigma_{\hat{y}} + C_2}{\sigma_y^2 + \sigma_{\hat{y}}^2 + C_2} \tag{2.3}$$

$$s(y, \hat{y}) = \frac{\sigma_{y\hat{y}} + C_3}{\sigma_y \sigma_{\hat{y}} + C_3} \tag{2.4}$$

where C_1, C_2, C_3 are small constants given by

$$C_1 = (K_1 L)^2, C_2 = (K_2 L)^2 \text{ and } C_3 = C_2/2, \tag{2.5}$$

respectively. L is the dynamic range of the pixel values (255 for 8-bit grayscale images), and $K_1 \ll 1, K_2 \ll 1$ are small scalar constants. In [16] K_1 and K_2 are set to 0.01 and 0.03, respectively. The general form of the Structural SIMilarity index between signal y and $\hat{y}$ is defined as:

$$\mathrm{SSIM}(y, \hat{y}) = [l(y, \hat{y})]^\alpha [c(y, \hat{y}]^\beta [s(y, \hat{y}]^\gamma, \tag{2.6}$$

where α, β and γ are parameters to define the relative importance of the three components. Specifically, we set $\alpha = \beta = \gamma = 1$, and the resulting SSIM index is given by

$$\mathrm{SSIM}(y,\hat{y}) = \frac{\left(2\mu_y\mu_{\hat{y}} + C_1\right)\left(2\sigma_{y\hat{y}} + C_2\right)}{\left(\mu_y^2 + \mu_{\hat{y}}^2 + C_1\right)\left(\sigma_y^2\sigma_{\hat{y}}^2 + C_2\right)} \tag{2.7}$$

which satisfies the following conditions:

1. symmetry: $\mathrm{SSIM}(y,\hat{y}) = \mathrm{SSIM}(\hat{y},y)$;
2. boundedness: $\mathrm{SSIM}(y,\hat{y}) \leqslant 1$;
3. unique maximum: $\mathrm{SSIM}(y,\hat{y}) = 1 \iff y = \hat{y}$.

In practise the SSIM index is applied locally, using an 11×11 circular-symmetric Gaussian weighting function $w = \{w_i | i = 1, 2, \ldots, N\}$, with standard deviation of 1.5 samples, normalized to unit sum $\left(\sum_{i=1}^{N} w_i = 1\right)$.

$$\mathrm{MSSIM}(y,\hat{y}) = \frac{1}{M}\sum_{j=1}^{M} \mathrm{SSIM}(y_j,\hat{y}_j) \tag{2.8}$$

where y_j and $\hat{y}_j$ are the image contents at the j-th local windows, and M is the number of local windows of the image.

The local statistics should be modified according to w:

- $\mu_y = \sum_{i=1}^{N} w_i y_i$
- $\sigma_y = \left(\sum_{i=1}^{N} w_i (y_i - \mu_y)^2\right)^{\frac{1}{2}}$
- $\sigma_{y\hat{y}} = \sum_{i=1}^{N} w_i (y_i - \mu_y)(\hat{y}_i - \mu_{\hat{y}})$.

2.2 MS-SSIM

A single-scale method as described in the previous section may be appropriate only for specific settings. Multi-scale method is a convenient way to incorporate image details at different resolutions. The authors of [17] proposed a multi-scale SSIM method that taking the reference and distorted image signals as the input, the system iteratively applies a low-pass filter and downsamples the filtered image by a factor of 2. The original image has been indexes as Scale 11, while the highest scale as Scale M, which is obtained after $M-1$ iterations. At the j-th scale, the contrast comparison 2.3 and the structure comparison 2.4 are calculated and denoted as $c_j(x, y)$ and $s_j(x, y)$, respectively. The luminance comparison 2.2 is computed only at Scale M and is denoted as $l_M(x, y)$. The overall SSIM evaluation is obtained by combining the measurement at different scales using

$$\mathrm{SSIM}(y, \hat{y}) = [l_M(y, \hat{y})]^{\alpha_M} \cdot \prod_{j=1}^{M} [c_j(y, \hat{y})]^{\beta_j} [s_j(y, \hat{y})]^{\gamma_j} \tag{2.9}$$

To calibrate the system parameters, the authors of [17] involved in thier experiments 8 subjects, asking them to assess the quality of synthesized images. With this approach they estimated $\beta_1 = \gamma_1 = 0.0448$, $\beta_2 = \gamma_2 = 0.2856$, $\beta_3 = \gamma_3 = 0.3001$, $\beta_4 = \gamma_4 = 0.2363$, and $\alpha_5 = \beta_5 = \gamma_5 = 0.1333$, respectively.

2.3 LPIPS-COMP

The same technique used to train LPIPS has been adopted for LPIPS-Comp. While LPIPS is trained on BAPPS, that contains images with several distortions but accounts only for compression artefacts from JPEG, LPIPS-Comp has seen a compression specific perceptual similarity dataset. In doing so, experiments showed that LPIPS-Comp aligns more to human judgement than the standard LPIPS on general compression tasks.

LPIPS-Comp [11] is a perceptual similarity metric based on deep neural networks obtained following the same approach as in [19] with LPIPS. These methods employ the ReLU activations after each *conv* block in the first five layers of the VGG-16 [14] architecture, with batch-normalization [4].

Feed-forward is performed on VGG-16 for both the original (y) and the reconstructed image ($\hat{y}$). Let L be the set of layers used for loss calculation (five for our setup), a function $F(y)$ denoting feed-forward on an input image y. $F(y)$ and $F(\hat{y})$ return two stacks of feature activation's for all L layers.

The LPIPS-Comp loss is then computed as:

- $F(y)$ and $F(\hat{y})$ are unit-normalized in the channel dimension. Let us call these, $z_y^l, z_{\hat{y}}^l \in R^{H_l \times W_l \times C_l}$ where $l \in L$. (H_l, W_l are the spatial dimensions).
- $z_y^l, z_{\hat{y}}^l$ are scaled channel wise by multiplying with the vector $w_l \in R^{C_l}$.

- The L_2 distance is then computed and an average over spatial dimension is taken. Finally, a channel-wise sum is performed.

Equation 2.10 summarizes the LPIPS-Comp computation.

$$\text{LPIPS-Comp}(y, \hat{y}) = \sum_{l} \frac{1}{H_l W_l} \sum_{h,w} \left\| w_l \odot \left(z^l_{\hat{y},h,w} - z^l_{y,h,w} \right) \right\|_2^2 \quad (2.10)$$

Note that the weights in F are learned for image classification on the ImageNet dataset [13] and are kept fixed. w are the linear weights learned on top of F on the BAPPS [19] dataset for LPIPS and on a compression specific similarity dataset for LPIPS-Comp. While in the first, compressed images are obtained with JPEG only, the second makes use of several compression methods: Mentzer et al. [9], Patel et al. [10], BPG [1] and JPEG-2000 [15].

2.4 DISTS

The authors of DISTS [] carried out five major experiments. First, they showed that DISTS has not the best performance overall on LIVE [], CSIQ [], and TID2013 [] that are datasets that have been around in the literature long enough to be likely overfitted by recent quality measures. Second, they saw comparable results on the BAPPS dataset against LPIPS (that is a metric trained on the BAPPS dataset). Third, DISTS achieves best performance...

2.5 SFSN

The authors of [21] found that a linear combination of a local structural fidelity assessment (SF) and a global statistical naturalness measure (SN) achieves high correlation with human judgement (measured with MOS on public Single Image Super Resolution IQA datasets, such as WIND [18], CVIU [8], and QADS [20]).

3

CONCLUSION

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

BIBLIOGRAFIA

[1] F. Bellard. Bpg image format. `https://bellard.org/bpg/`, 2014. (Cited on page 13.)

[2] K. Ding, K. Ma, S. Wang, and E. P. Simoncelli. Image quality assessment: Unifying structure and texture similarity. *arXiv preprint arXiv:2004.07728*, 2020. (Cited on page 9.)

[3] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30, 2017. (Cited on page 9.)

[4] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pages 448–456. PMLR, 2015. (Cited on page 12.)

[5] M. Kettunen, E. Härkönen, and J. Lehtinen. E-lpips: robust perceptual image similarity via random transformation ensembles. *arXiv preprint arXiv:1906.03973*, 2019. (Cited on page 9.)

[6] A. Kirillova, E. Lyapustin, A. Antsiferova, and D. Vatolin. Erqa: Edge-restoration quality assessment for video super-resolution. *arXiv preprint arXiv:2110.09992*, 2021. (Cited on page 9.)

[7] E. Lyapustin, A. Kirillova, V. Meshchaninov, E. Zimin, N. Karetin, and D. Vatolin. Towards true detail restoration for super-resolution: A benchmark and a quality metric. *arXiv preprint arXiv:2203.08923*, 2022. (Cited on page 9.)

[8] C. Ma, C.-Y. Yang, X. Yang, and M.-H. Yang. Learning a no-reference quality metric for single-image super-resolution. *Computer Vision and Image Understanding*, 158:1–16, 2017. (Cited on page 13.)

[9] F. Mentzer, E. Agustsson, M. Tschannen, R. Timofte, and L. Van Gool. Conditional probability models for deep image compression. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4394–4402, 2018. (Cited on page 13.)

[10] Y. Patel, S. Appalaraju, and R. Manmatha. Deep perceptual compression. *arXiv preprint arXiv:1907.08310*, 2019. (Cited on page 13.)

[11] Y. Patel, S. Appalaraju, and R. Manmatha. Saliency driven perceptual image compression. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 227–236, 2021. (Cited on pages 9 and 12.)

[12] M. Rastegari, V. Ordonez, J. Redmon, and A. Farhadi. Xnor-net: Imagenet classification using binary convolutional neural networks. In *European conference on computer vision*, pages 525–542. Springer, 2016.

[13] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al. Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3):211–252, 2015. (Cited on page 13.)

[14] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. (Cited on page 12.)

[15] A. Skodras, C. Christopoulos, and T. Ebrahimi. The jpeg 2000 still image compression standard. *IEEE Signal processing magazine*, 18(5):36–58, 2001. (Cited on page 13.)

[16] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004. (Cited on pages 9 and 10.)

[17] Z. Wang, E. P. Simoncelli, and A. C. Bovik. Multiscale structural similarity for image quality assessment. In *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, volume 2, pages 1398–1402. Ieee, 2003. (Cited on pages 9, 11, and 12.)

[18] H. Yeganeh, M. Rostami, and Z. Wang. Objective quality assessment of interpolated natural images. *IEEE Transactions on Image Processing*, 24(11):4651–4663, 2015. (Cited on page 13.)

[19] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang. The unreasonable effectiveness of deep features as a perceptual metric.

In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 586–595, 2018. (Cited on pages 9, 12, and 13.)

[20] F. Zhou, R. Yao, B. Liu, and G. Qiu. Visual quality assessment for super-resolved images: Database and method. *IEEE Transactions on Image Processing*, 28(7):3528–3541, 2019. (Cited on page 13.)

[21] W. Zhou, Z. Wang, and Z. Chen. Image super-resolution quality assessment: Structural fidelity versus statistical naturalness. In *2021 13th International Conference on Quality of Multimedia Experience (QoMEX)*, pages 61–64. IEEE, 2021. (Cited on pages 9 and 13.)