SINGLE IMAGE RESTORATION WITH GENERATIVE PRIORS

Kalliopi Basioti

Rutgers University, USA
kalliopi.basioti@rutgers.edu

George V. Moustakides

University of Patras, Greece
moustaki@upatras.gr

2021 IEEE International Conference on Image Processing (ICIP) | 978-1-6654-4115-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICIP42928.2021.9506768

ABSTRACT

Generative models can be used, as an alternative to conventional probability densities, to capture the statistical behavior of complicated datasets. Unlike probability densities with which the generation of realizations may become a challenging task, generative models have an inherent ability to easily produce realizations, which, in the case of natural images can be extremely realistic. In many image restoration problems, such as deblurring, colorization, inpainting, super-resolution, etc., probability densities are used as priors, one may therefore wonder whether we can, instead, adopt generative models. Indeed such methods have appeared in the literature, but they require exact knowledge of the transformations responsible for the data distortion and involve regularizer terms with weights that require adjustment. Our approach, by combining maximum a-posteriori probability with maximum likelihood estimation, can successfully restore images in both blind and non-blind modes without the need to fine-tune any regularization parameters. Simulations on deblurring, colorization, and image separation problems with exact knowledge of the transformation demonstrate improved image quality, reduced computational cost compared to existing methods. Comparable results are also enjoyed when the distortion models contain unknown parameters.

Index Terms— Image restoration, Image separation, Blind image restoration/separation, Bayes procedures, Generative modeling

1. INTRODUCTION

A standard mathematical model for the image restoration problem is given by

$$Y = T(X, \alpha) + W, \tag{1}$$

where X is a hidden vector, representing the original image; $T(X, \alpha)$ is a deterministic transformation with known functional form that can possibly contain unknown parameters $\alpha \in \mathcal{A}$, with $\mathcal{A}$ a known set; W is a random vector independent from X that expresses additive noise and/or modeling error. For W we assume that it is distributed according to the density $g(W, \beta)$ which has a known functional form and possibly unknown parameters $\beta \in \mathcal{B}$, with $\mathcal{B}$ a known set.

The image restoration problem is known for its ill-posedness. Therefore, for its solution, prior knowledge for the original data must be incorporated. Traditional examples of image priors use Student's t-distribution for noise modeling [1], or spatially varying priors [2] or total-variation [3] for more difficult problems as inpainting. Although these approaches perform well in restoration problems such as denoising and deblurring, they tend to fail to reconstruct image details in applications such as super-resolution, inpainting, or colorization, suggesting the need for more descriptive priors.

Generative models can be successfully trained to generate realizations from an unknown distribution for which we have available training data. Classical example constitute the Generative Adversarial Networks (GANs) [4], which are known to create incredibly realistic natural images [5]. Because of this fact, several approaches attempted to solve inverse problems using generative priors for the original data. Early efforts were using the generative model partially (only the generator function) [6, 7, 8, 9]. Only recently [10, 11] we see techniques that employ the complete model (generator and input density) improving the performance of the corresponding methods. However, a drawback of these methods is the existence of unknown weighting parameters. As a result, they require data pairs of original and transformed data used in additional simulations to find the appropriate values for their parameters. Our methodology, which relies on statistical estimation theory, can identify all parameters entering the formulation, mitigating the need to calibrate unknown quantities. Moreover, the method we will develop will treat cases where the transformation responsible for the image distortion is not precisely known as required by all existing techniques. Finally, we must point out that our approach can address restoration problems in dynamic environments where the unknown image degradation mechanism may change with every single realization (image) of X.

2. RESTORATION WITH GENERATIVE MODELING

A generative model for a random vector X consists of a transformation $G(Z)$, known as the generator function, and an input density $h(Z)$. We can then generate realizations of X by generating realizations of Z that follow $h(Z)$ and then trans-

This work was supported by the US National Science Foundation under Grant CIF 1513373, through Rutgers University.

forming them $X = G(Z)$ through the generator. In image restoration we are interested in the recovery of X from the observation Y when X and Y are related through (1). To obtain such an estimate, we intend to exploit the generative model in the following way: Instead of finding the estimate $\hat{X}$ directly as in [10, 11], we propose to obtain $\hat{Z}$ of the input to the generator and then estimate X as $\hat{X} = G(\hat{Z})$. We note that with this approach we can accommodate generator functions that are not necessarily invertible, suggesting that X may even lie on a lower dimensional manifold.

Let us first produce the joint density of Y, Z given α, β. We have

$$f(Y, Z|\alpha, \beta) = g\big(Y - T(G(Z), \alpha)|\beta\big)h(Z). \quad (2)$$

To estimate Z we intend to apply the MAP estimator. Of course there is the problem of the unknown parameters α, β. In Statistical estimation theory, unknown parameters (quantities for which we have no prior density) can be estimated using the maximum likelihood estimator. Consequently, we propose the simultaneous estimation of Z, α, β which gives rise to the following optimization problem

$$\hat{Z} = \arg\max_{Z}\Big(\max_{\alpha\in\mathcal{A},\beta\in\mathcal{B}} g\big(Y - T(G(Z), \alpha)|\beta\big)\Big)h(Z). \quad (3)$$

Detailed mathematical derivation of this equation based on classical parameter estimation theory can be found in [12].

2.1. Gaussian Noise and Gaussian Input

Let us now specify in more detail our mathematical model. For the additive noise vector W appearing in (1), we assume that it has Gaussian elements independent and identically distributed with mean zero and variance β^2 (we adopt the Gaussian model only for simplicity), namely $g(W|\beta)$ is $\mathcal{N}(0, \beta^2 I)$ where I is identity matrix. We then conclude that

$$\max_{\beta\geq 0} g(W|\beta) = \frac{C}{\|W\|^N}, \quad (4)$$

where C constant and N is the size of the vector W. If we also select the input density $h(Z)$ to be Gaussian $\mathcal{N}(0, I)$ then, for a transformation $T(X)$ without unknown parameters, we can write for (3) that the estimates are equivalent to

$$\hat{Z} = \arg\min_{Z}\big\{\|Z\|^2 + N\log\|Y - T\big(G(Z)\big)\|^2\big\}. \quad (5)$$

We can now compare our optimization problem in (5) with the most efficient existing techniques that are estimating Z by solving the problem

$$\hat{Z} = \arg\min_{Z}\big\{\|Y - T\big(G(Z)\big)\|^2 + \lambda\|Z\|^{\nu}\big\} \quad (6)$$

where $\nu = 1$ in [11], $\nu = 2$ in [7, 10]. As we can see in our approach there is no weighting parameter λ therefore, no prior fine-tuning is necessary. A notable difference is also how the error distance $\|Y - T\big(G(Z)\big)\|^2$ is combined with the input power $\|Z\|^2$. In our method we use the logarithm of the distance while in [7, 10, 11] it is the distance itself combined with $\|Z\|$ or $\|Z\|^2$.

2.2. Parametric Transformations

Let us focus on the more challenging problem of a transformation $T(X, \alpha)$ containing unknown parameters α. Following our general theory developed for the case of noise and generator input being Gaussian, the MAP estimator with maximum likelihood estimation of the parameters takes the form

$$\hat{Z} = \arg\min_{Z}\big\{N\log\big(\min_{\alpha}\|Y - T(G(Z), \alpha)\|^2\big) + \|Z\|^2\big\}. \quad (7)$$

This general version of the problem is equivalent to

$$\{\hat{Z}, \hat{\alpha}\} = \arg\min_{Z,\alpha}\big\{N\log\|Y - T(G(Z), \alpha)\|^2 + \|Z\|^2\big\} \quad (8)$$

and the joint minimization can be carried out, for example, with a simple steepest descent algorithm.

To further advance our analysis consider linear transformations of the form $T(X, \alpha) = T(\alpha)X$, where $T(\alpha)$ is a matrix. In fact these are the most common transformations in image restoration problems. Furthermore, assume that $T(\alpha)$ can be decomposed as:

$$T(\alpha) = \alpha_1 T_1 + ... + \alpha_M T_M \quad (9)$$

where $T_1, ..., T_M$ are known matrices and $\alpha = [\alpha_1, ..., \alpha_M]^{\intercal}$ is the unknown parameter vector. For this case it is possible to obtain a more convenient expression for the optimization problem. Define $\mathcal{T} = [T_1 G(Z), ..., T_M G(Z)]$ and concentrate on the minimization over α in (7). Using (9) we can perform it analytically. Indeed by focusing on the distance inside the logarithm we observe that

$$\begin{aligned}\min_{\alpha}\|Y - T(\alpha)G(Z)\|^2 &= \min_{\alpha}\|Y - \mathcal{T}\alpha\| \\ &= \|Y\|^2 - Y^{\intercal}\mathcal{T}(\mathcal{T}^{\intercal}\mathcal{T})^{-1}\mathcal{T}^{\intercal}Y \quad (10)\end{aligned}$$

with the last outcome following from the Orthogonality Principle, see [13], pp.79-83, and expressing the projection error of Y onto the linear subspace spanned by the columns of $\mathcal{T}$. This result, when substituted in (7), yields

$$\hat{Z} = \arg\min_{Z}\big\{\|Z\|^2 + N\log(\|Y\|^2 - Y^{\intercal}\mathcal{T}(\mathcal{T}^{\intercal}\mathcal{T})^{-1}\mathcal{T}^{\intercal}Y)\big\} \quad (11)$$

where the only minimization is with respect to Z and where, as we recall from its definition, $\mathcal{T}$ is a matrix that depends on Z as well.

2.3. The data separation problem

In [14, 15, 16, 17], existing single data vector methods extend to mixtures of multiple random vectors where a separate generative model describes each participating vector. These extensions experience the same drawbacks as their original single vector counterparts: (1) They contain multiple regularizer terms with unknown weights that need to be tuned appropri-

ately; (2) The corresponding methods cannot accommodate mixtures involving unknown parameters. And in the case of [17], they are also restricted to likelihood-based generative models.

We can overcome the previous limitations by generalizing our technique to cover combinations of multiple vectors. For simplicity we only treat the two vector case with the extension to any number of vectors being straightforward. Suppose that we have two random vectors X_1, X_2 each satisfying a generative model $X_i = G(Z_i)$ with input density $Z_i \sim h_i(Z_i), i = 1, 2$. If the mixture data model satisfies

$$Y = \alpha_1 X_1 + \alpha_2 X_2 + W \tag{12}$$

where the additive noise/modeling error W has density $g(W|\beta)$ with parameters β then we can combine all parts and produce the joint probability density

$$f(Y, Z_1, Z_2|\alpha_1, \alpha_2, \beta) = \\ g(Y - \alpha_1 G_1(Z_1) - \alpha_2 G_2(Z_2)|\beta) h_1(Z_1) h_2(Z_2). \tag{13}$$

In (13) we silently made the assumption that Z_1, Z_2 (and therefore X_1, X_2) are statistically independent which produces the product of the two input densities. Following our general methodology we need to perform the following optimization

$$\max_{Z_1, Z_2} \max_{\alpha_1, \alpha_2, \beta} f(Y, Z_1, Z_2|\alpha_1, \alpha_2, \beta). \tag{14}$$

If, as before, $g(W|\beta)$ is Gaussian with mean 0 and covariance $\beta^2 I$ and both input vectors are independent Gaussian with mean 0 and unit covariance matrix then the previous maximization after first maximizing over β is equivalent to

$$\{\hat{Z}_1, \hat{Z}_2\} = \arg \min_{Z_1, Z_2} \Big\{ \|Z_1\|^2 + \|Z_2\|^2 + \\ N \log \big(\min_{\alpha_1, \alpha_2} \|Y - \alpha_1 G(Z_1) - \alpha_2 G(Z_2)\|^2 \big) \Big\}. \tag{15}$$

We can either apply gradient descent on the combination $\{Z_1, Z_2, \alpha_1, \alpha_2\}$ or solve analytically for $\{\alpha_1, \alpha_2\}$, substitute, and then minimize over $\{Z_1, Z_2\}$. Regarding the latter, thanks to the Orthogonality Principle we can write

$$\min_{\alpha_1, \alpha_2} \|Y - \alpha_1 G_1(Z_1) - \alpha_2 G_2(Z_2)\|^2 = \\ \|Y\|^2 - Y^\intercal \mathcal{G} (\mathcal{G}^\intercal \mathcal{G})^{-1} \mathcal{G}^\intercal Y, \tag{16}$$

where $\mathcal{G} = [G_1(Z_1), G_2(Z_2)]$. Substituting in (15) yields

$$\{\hat{Z}_1, \hat{Z}_2\} = \arg \min_{Z_1, Z_2} \Big\{ \|Z_1\|^2 + \|Z_2\|^2 + \\ N \log \big(\|Y\|^2 - Y^T \mathcal{G} (\mathcal{G}^T \mathcal{G})^{-1} \mathcal{G}^T Y \big) \Big\}, \tag{17}$$

where we recall that $\mathcal{G}$ is a matrix that depends on $\{Z_1, Z_2\}$ as well.

Similarly, it is possible to accommodate the more general case of a nonlinear mixing function $T(X_1, X_2, \alpha)$ with possibly unknown parameters α, which combines the two vectors and generates the single observation vector Y through $Y = T(X_1, X_2, \alpha) + W$. Under the Gaussian assumption for the two inputs and the additive noise term, we can obtain the estimates $\hat{Z}_1, \hat{Z}_2$ by solving the optimization problem

$$\{\hat{Z}_1, \hat{Z}_2, \hat{\alpha}\} = \arg \min_{Z_1, Z_2, \alpha} \Big\{ \|Z_1\|^2 + \|Z_2\|^2 + \\ N \log \big(\|Y - T\big(G_1(Z_1), G_2(Z_2), \alpha\big)\|^2 \big) \Big\}. \tag{18}$$

Of course with general nonlinear mixers we can no longer solve for α and substitute, as in the linear case. Consequently the minimization must be carried out simultaneously for the triplet $\{Z_1, Z_2, \alpha\}$. Finally, as in the single vector case, restorations are defined as $\hat{X}_i = G_i(\hat{Z}_i), i = 1, 2$.

3. EXPERIMENTS

For our experiments, we use the CelebA [18] and the Caltech-UCSD Birds [19] datasets. The first dataset contains 202,599 RGB images cropped and resized to 64x64x3 and then separated into two sets of 202,499 for training and 100 for testing. For the Birds dataset, we train two models, one with the original images and the second with segmented images, with a removed background using the included segmentation masks. In both cases, the images are resized to $64 \times 64 \times 3$ while we kept 10,609 images for training and 1179 for testing. We trained a progressive, growing GAN for each of cases using the training sets, as described in [5] for 64×64 images. Finally, we applied the momentum gradient descent [20] with normalized gradients in all competing methods, with the momentum hyperparameter set to 0.999 and the learning rate to 0.001. We compare the different methods in terms of the PSNR and structural similarity index measure (SSIM) [21].

In our first experiment, we investigate the image deblurring problem. We blur the original images from CelebA with the standard 3×3 Gaussian kernel. For our second set of simulations, we recover an RGB image from one of its chromatic components. As such, we select the green channel. For the two problems, we compare the methods of Yeh et. al. [6] and Bora et. al, Whang et. al., Asim et. al. [7, 10, 11] (the two approaches coincide for gaussian noise) with our method. The techniques in [6, 7, 10, 11] require exact knowledge of the kernel coefficients. They also need fine-tuning of their regularization parameter weights, which is achieved by solving multiple instances of their optimization problem with various weight values and selecting the one delivering the smallest average reconstruction error. For [6] the best calibrated weight value was 0.6 (for Gaussian deblurring) and 0.1 (for colorization), while for [7, 10, 11] 0.6 for Gaussian deblurring and 0.5 for colorization.

Since our method has no unknown weights, no tuning phase is necessary. We distinguish two versions of our approach. In the first, we assume that we know the kernel (blur-

Table 1. PSNR and SSIM scores

	Deblurring		Colorization	
	PSNR	SSIM	PSNR	SSIM
[6]	27.28	0.914	21.83	0.789
[7, 10, 11]	28.28	0.925	21.35	0.779
Proposed	28.20	0.922	21.85	0.790
Proposed^{B}	26.95	0.901	21.31	0.805

Table 2. PSNR and SSIM scores

	$\alpha_1 = \alpha_2 = 0.5$				$\alpha_1 = \alpha_2 = 0.5$			
	Faces		Birds Seg.		Faces		Birds	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
[16]	25.23	0.821	22.74	0.831	18.62	0.678	18.57	0.738
Proposed	27.21	0.840	25.47	0.906	19.23	0.718	19.03	0.800
Proposed^{B}	27.00	0.846	24.50	0.902	18.49	0.702	18.33	0.753

Fig. 1. For the first two sets of images: Column a) Original; b) Transformed; c) [6], known parameters; d) [7, 10, 11], known parameters; e) Proposed, known parameters; f) Proposed, unknown parameters. In the first set, deblurring of a Gaussian kernel. In the second set, colorization of a monochromatic image. For the second two sets of images: Columns a), b) Originals; c) Mixture; d), e) [16], known coefficients; f), g) Proposed, known coefficients; h), i) Proposed, unknown coefficients. In both sets, image separation. More experiments and comparisons can be found in [12, 22].

ring) coefficients and also the selected channel to make fair comparisons with [6], [7, 10, 11]. In the second version, we assume that the blurring coefficients and the selected channel are unknown, implying that we simultaneously estimate them and restore the original image. For the deblurring, we solve (8) under the constraint that the filter coefficients sum to one. For the colorization, we notice that the channel decomposition is a linear transformation implemented with three matrices $T_{\mathrm{R}}, T_{\mathrm{G}}, T_{\mathrm{B}}$ as in (10). The fact that the unknown parameter is now discrete does not pose any special difficulty in the optimization in (8), which must be modified as follows

$$\hat{Z} = \arg\min_{Z} \left\{ \|Z\|^2 + N \log \left(\min_{i=\mathrm{R,G,B}} \|Y - T_i \mathcal{G}(Z)\|^2 \right) \right\}.$$

The first two sets of images in Fig. 1 show four examples of image deblurring and colorization from left to right. We also see in Table 1 the corresponding PSNRs and SSIMs. We realize that the proposed methodology enjoys comparable restoration quality as the existing methods for both deblurring and colorization when the transformations are known. However, in our case, this is achieved without the computational overhead of weight fine-tuning and the need for original-transformed pairs of images. Furthermore, in both scenarios, it delivers similar quality if the parameters of the transformation are considered unknown and need to be estimated in parallel with the original image.

In our last set of experiments, we create linear mixtures of CelebA and Caltech-UCSD Birds. We begin by separating faces from the segmented Caltech-UCSD birds with mixture coefficients $\alpha_1 = \alpha_2 = 0.5$. Next, we switch to the full Caltech-UCSD dataset with $\alpha_1 = \alpha_2 = 0.5$. We mention that for all three techniques, the evaluated methods are identical. As in single image restoration, we distinguish two versions of our methodology, namely the solution of (15) with α_1, α_2 known and the solution of 18 when the two parameters are considered unknown. We compare our two versions against the results obtained by solving

$$\{\hat{Z}_1, \hat{Z}_2\} = \arg\min_{Z_1, Z_2} \Big\{ \|Y - \alpha_1 G_1(Z_1) - \alpha_2 G_2(Z_2)\|^2 + \lambda_1 \|Z_1\|^2 + \lambda_2 \|Z_2\|^2 \Big\},$$

which is the method proposed in Soltani et al. [16] and contains two weighting parameters λ_1, λ_2 that require tuning, and exact knowledge of α_1, α_2. Extending the tuning method of the single weight to two weights, we obtained the calibrated values $\lambda_1 = \lambda_2 = 0.3$, for the mixture with segmented birds, and $\lambda_1 = 0.5, \lambda_2 = 0.4$ for the mixtures with the original birds. In the third and fourth set of images in Fig. 1, we present examples of the two mixture types and the corresponding results of the competing separation methods, while in Table 2 we give the corresponding PSNRs and SSIMs per dataset. We observe that our method improves the quality of the separated images without requiring parameter fine-tuning and knowledge of the mixture parameters.

4. CONCLUSION

We introduced a general restoration methodology based on a generative model description of the class of original data. Our approach can successfully restore images through a well-defined mathematical optimization problem that does not require any fine-tuning of weights of regularizer terms, which is standard in existing methods. Our method's most notable advantage is its ability to restore data even when the transformation responsible for their deformation contains unknown parameters. Experiments using popular image datasets show that our method can deliver similar restoration quality as the existing state of the art without exact knowledge of the transformation and the need for weight tuning of regularizer terms.

5. REFERENCES

[1] G. Chantas, N. Galatsanos, A. Likas, and M. Saunders, "Variational bayesian image restoration based on a product of t-distributions image prior," *IEEE Transactions on Image Processing*, vol. 17, no. 10, pp. 1795–1805, 2008.

[2] J. Mateos, T.E. Bishop, R. Molina, and A.K. Katsaggelos, "Local bayesian image restoration using variational methods and gamma-normal distributions," in 16th IEEE International Conference on Image Processing (ICIP), pp. 129–132, 2009.

[3] M.V. Afonso, J.M. Bioucas-Dias, and M. Figueiredo, "An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 681–695, 2010.

[4] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[5] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of gans for improved quality, stability, and variation," *arXiv:1710.10196*, 2017.

[6] R.A. Yeh, C. Chen, T. Yian Lim, A.G. Schwing, M. Hasegawa-Johnson, and M.N. Do, "Semantic image inpainting with deep generative models," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5485–5493, 2017.

[7] A. Bora, A. Jalal, E. Price, and A.G. Dimakis, "Compressed sensing using generative models," *arXiv:1703.03208*, 2017.

[8] R.A. Yeh, T.Y. Lim, C. Chen, A.G. Schwing, M. Hasegawa-Johnson, and M.N. Do, "Image restoration with deep generative models," in *EEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6772–6776, 2018.

[9] P.R. Siavelis, N. Lamprinou, and E.Z. Psarakis, "An improved gan semantic image inpainting," in *International Conference on Advanced Concepts for Intelligent Vision Systems*. Springer, 2020, pp. 443–454.

[10] J. Whang, Q. Lei, and A.G. Dimakis, "Compressed sensing with invertible generative models and dependent noise," *arXiv:2003.08089*, 2020.

[11] M. Asim, M. Daniels, O. Leong, A. Ahmed, and P. Hand, "Invertible generative models for inverse problems: mitigating representation error and dataset bias," in *International Conference on Machine Learning*. PMLR, pp. 399–409, 2020.

[12] K. Basioti and G.V. Moustakides, "Image restoration from parametric transformations using generative models," *arXiv: 2005.14036*, 2020.

[13] B. Hajek, *Random Processes for Engineers*, Cambridge University Press, 2015.

[14] R. Anirudh, J.J. Thiagarajan, B. Kailkhura, and T. Bremer, "An unsupervised approach to solving inverse problems using generative adversarial networks," *arXiv:1805.07281*, 2018.

[15] Y.C. Subakan and P. Smaragdis, "Generative adversarial source separation," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 26–30, 2018.

[16] M. Soltani, S. Jain, A.V. Sambasivan, and C. Hegde, "Leaming structured signals using gan s with applications in denoising and demixing," in *53rd Asilomar Conference on Signals, Systems, and Computers*, pp. 2127–2131, 2019.

[17] V. Jayaram and J. Thickstun, "Source separation with deep generative priors," *arXiv:2002.07942*, 2020.

[18] Z. Liu, P. Luo, X. Wang, and X. Tang, "Deep learning face attributes in the wild," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3730–3738, 2015.

[19] C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie, "The CalTech-UCSD Birds-200-2011 dataset," 2011.

[20] N. Qian, "On the momentum term in gradient descent learning algorithms," *Neural Networks*, vol. 12, no. 1, pp. 145–151, 1999.

[21] A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *20th International Conference on Pattern Recognition*, pp. 2366–2369, 2010.

[22] K. Basioti and G.V. Moustakides, "Image dequantization using generative models as priors," *arXiv: 2007.07923*, 2020.