A FRAMEWORK FOR CONTRAST ENHANCEMENT ALGORITHMS OPTIMIZATION

Simone Zini, Marco Buzzelli, Simone Bianco, Raimondo Schettini

Department of Informatics, Systems and Communication
University of Milano - Bicocca
Viale Sarca 336, Milan 20126, Italy

ABSTRACT

We present a general-purpose framework for the optimization of parametric contrast enhancement algorithms. We first define a regression module for image acceptability, which is based on deep neural features and which is trained on a large dataset of user-expressed preferences. This regression module is then used as the objective function of a Bayesian optimization process, guiding the search for the optimal parameters of a given contrast enhancement algorithm. In our experiments we optimize three different contrast enhancement algorithms of varying levels of complexity. The effectiveness of our optimization framework is experimentally confirmed by evaluating the output of the optimized contrast enhancement algorithms with respect to reference enhanced images.

***Index Terms*—** Contrast enhancement, Bayesian Optimization, Convolutional Neural Networks

1. INTRODUCTION

Image contrast enhancement is one of the last operations conducted in a typical camera pipeline, and it serves the purpose of improving the perceivability of objects in the scene by enhancing the brightness difference between objects and their backgrounds [1].

The problem has been treated with different approaches in the state of the art, from the simplest gamma correction and histogram equalization, up to transform based methods, exposure-based methods and image fusion based methods. Most of the existing approaches for image contrast enhancement rely on the values of one or more parameters to operate on images in order to perform the correction steps. These parameters are in general tuned manually on a set of different possible case scenarios or for specific images.

In this paper a user-preferences based framework for contrast enhancement algorithms optimization is proposed, as shown in Figure 1. The proposed framework is based on the use of a logistic regressor, capable to model user preferences based on the concept of image acceptability defined by

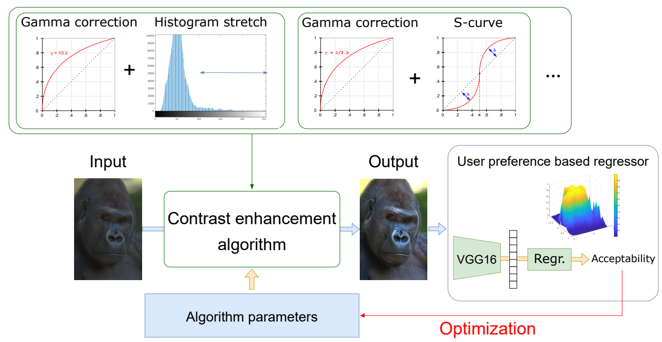

Fig. 1. Overview of the proposed framework for contrast enhancement algorithms optimization.

Jaroensri et al. [2] and described later on. The logistic regressor score given to new images is used as objective function for bayesian optimization for the selection of the best parameters of different algorithms for image contrast enhancement.

Several methods for adjusting image contrast have been developed through the years. In general two groups of contrast enhancement algorithms can be identified: based on global correction or local correction. In the first group is possible to find approaches like gamma correction and histogram equalization. Multiple versions of the histogram equalization technique have been proposed in the years, making it adaptive to the content of the image [3], preserving the original image brightness while enhancing the contrast [4, 5], or incorporating models of perception [6, 7]. Other approaches in this first group are exposure-based methods [8], which adjust the exposure level of an image using a mapping function between the light values and the pixel values of interested objects, and image fusion based methods [9], which combine relevant information from multiple images taken from the same scene in order to produce a final more informative one. In the second group, related to local correction, it is possible to find the method by Moroney et al. [10], and the Local Contrast Correction (LCC) by Schettini et al. [11]: these are based on either the Gaussian or the Bilateral filter to selectively determine image areas to be brightened or darkened. In the last years a number of approaches for image contrast enhance-

Source code and data available at: `https://github.com/TheZino/contrast-algo-opt-framework` ment exploiting machine learning have also been presented. Here can be found solutions based on Neural Networks for image enhancement [12] and techniques for hyperparameter selection and optimization of specific algorithms [13, 14]. Jaroensri et al. [2] proposed in 2015 a solution to model user-expressed preferences of image acceptability: starting from a subset of the MIT-Adobe 5K dataset [15], the authors modified the images using different combinations of contrast and brightness, which were later evaluated by a pool of users as being acceptable or not acceptable. A logistic regression classifier was then trained to fit the acceptability label based on low level image content features such as the fraction of highlight and shadow clipping, luminance histogram, root-mean-square luminance contrast, and others. In our work we develop an alterntive deep-feature-based acceptability regressor, which we exploit to guide an optimization process for image contrast enhancement.

2. PROPOSED FRAMEWORK FOR CONTRAST ENHANCEMENT

The proposed framework for contrast enhancement trains a logistic regressor for image acceptability estimation, and uses it as objective function to guide a Bayesian optimization for the selection parameters in contrast enhancement algorithms. A general overview of the approach is depicted in Figure 1.

2.1. Regression of user-expressed image acceptability

A logistic regressor is first trained with the purpose of using it as objective function of the contrast enhancement parameters optimization step, described in the next section. Where preliminary works relied on low-level image features for this task [2], here we adopt an approach based on convolutional neural networks, allowing for a much faster training procedure and inference step. Specifically, we rely on a VGG-16 model [16] previously trained for image classification on 1000 classes, extracting features from the last convolutional layer, after which an average pooling operation is applied in order to bring the spatial resolution to dimension 1×1. The features obtained from the deep neural network are used to train a tree ensemble regressor based on Adaptive logistic regression (LogitBoost) [17], minimizing binomial deviance. Different instances of the regressor are trained in relation to the data preprocessing as later described in section 3.1.

2.2. Parametric contrast enhancement

Given a model of user preferences capable to associate a score of acceptability to an image, is possible to design an optimization procedure in order to maximize the score of an image, given a certain enhancement algorithm. The proposed approach for algorithm optimization can be applied to any kind of algorithm for contrast enhancement whose performance depends on one or more parameters. In order to prove the effectiveness of user-preference-driven optimization, three algorithms for contrast enhancement are here considered.

The first algorithm consists of a simple combination of two global operators: first the image is processed using a gamma function (with parameter γ), and then an histogram stretching operation (with two parameters max_value and min_value) is performed over the output of the gamma function. The second one is a slightly different configuration which adopts a parametric S-curve function defined by Kang et al. [13], dependent on two parameters: λ, which determines the slope of the S-curve, and a which determines the flex point of the S-curve:

$$y = \begin{cases} a - a(1 - \frac{x}{a})^{\lambda} & if x \leq a \\ a + (1-a)(\frac{x-a}{1-a})^{\lambda} & otherwise \end{cases} \quad (1)$$

The last algorithm is Local Contrast Correction (LCC) by Schettini et al. [11], which determines areas of images to be brightened or darkened based on their local intensity, relying on a bilateral filter to reduce halo-like artifacts. The parameters of this algorithm are α, which determines the exponent of the gamma-like function applied to the luminance channel of the input images, and two standard deviation values σ_1 and σ_2 which are the parameters of the bilateral filter function. This algorithm works on the Y channel of the YCbCr color space.

2.3. Bayesian parameters optimization

We exploit Bayesian optimization [18] to determine the best parameters of the considered contrast enhancement algorithms. Let x be the N-dimensional parameter configuration of a given algorithm for contrast enhancement. Let $f(x)$, the objective function, be composed as: I) the application of the algorithm with parameters x on an input image or dataset, II) the extraction of VGG-16-based neural features from the enhanced images, III) the processing of neural features by the trained regressor, which assigns an acceptability score between 0 and 1. Let a be an acquisition function that evaluates the expected amount of improvement in the objective function. Bayesian optimization updates a Gaussian model of $f(x)$ to obtain a posterior distribution Q over functions, and finds the new point x that maximizes $a(x)$. In other words, it finds the set of parameters that maximizes the score given by the logistic regressor on the newly-enhanced images.

Two versions of each algorithm are considered for optimization: one by optimizing over a training dataset and one optimizing per individual image. In the first case the optimization is performed offline, optimizing the parameters based on the performance on the whole dataset, and then using these parameters for the corresponding algorithm at inference time. In the second case the optimization is performed directly over the images at processing time, thus leading to a set of parameters specifically optimized for each new image.

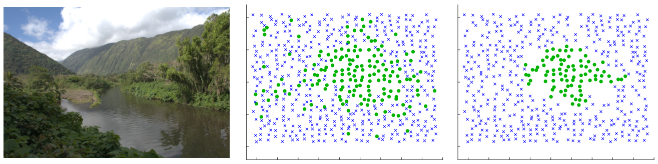

Fig. 2. Example of data points distribution in contrast/brightness space, before (left) and after (right) data cleaning. Green dots are images labelled as "acceptable" label, blue crosses are images labelled as "non acceptable".

3. EXPERIMENTAL SETUP

3.1. Acceptability dataset and data preprocessing

The dataset presented by Jaroensri et al. [2] has been adopted to train the logistic regressor of user-expressed acceptability. It contains user-expressed binary acceptability judgments of 500 images, adjusted to various configurations using brightness and contrast settings in roughly 600 variations. The images were originally taken from the MIT-Adobe fiveK Dataset [15], all processed with white balancing and saturation according to "Expert C". The total amount of data points collected is 301320, of which 241148 constitute the training set while the remaining 60174 are used as test set.

Due to the presence of an high amount of outliers, a cleaning procedure has been performed: as can be seen in Figure 2, in the areas where most of the images are labeled as being acceptable, sparse points with the non-acceptable label can occur, and viceversa. A visual inspection of these outliers confirmed them as being unequivocally mislabelled by the original users (such as completely dark images). In order to properly train the regressor model, those points have been removed from the training dataset. The cleaning procedure consists of a density analysis in the brightness/contrast space, performed by dividing the space in bins and counting the amount of positive labels in each bin. Using a fixed threshold the points have then been removed obtaining a new set of data points for each image. An example is shown in Figure 2.

Secondly, since in the original dataset the 600 data points per image present an high disparity in the labels (around 75% of the data points are labeled as not-acceptable, while the remaining 25% are labeled as acceptable), a data replication procedure has been applied to re-balance the training data distribution: the acceptable images have been replicated in order to reach the same amount of not-acceptable images.

3.2. Bayesian optimization dataset and configuration

The optimization procedure of the three contrast enhancement algorithms has been performed using the test set of the dataset proposed by Jaroensri et al. [2]. For each image in the dataset, the version with no contrast or brightness modification is used as input to the contrast-enhancement algorithm.

Table 1. Evaluation of user-expressed acceptability regression, analyzing the impact of the data augmentation and datapoint cleaning procedure on the dataset by [2].

Data balancing	Outliers removed	Micro accuracy	Macro accuracy	Precision	Recall	F-Score
		78.85%	64.26%	0.607	0.360	0.452
✓		75.01%	73.19%	0.489	0.697	0.575
✓	✓	74.98%	73.33%	0.488	0.701	0.576

Starting from this image and a set of random values for the algorithm parameters, each of the three algorithms have been optimized using the result of the logistic regressor as objective function. The optimization has been performed using Bayesian optimization procedure for a total amount of 30 iterations. In the case of optimization on training dataset, at each iteration 120 random images are processed and evaluated. In the case of per-image optimization, the procedure processed the same image for a total amount of 30 iterations.

3.3. Optimized contrast enhancement evaluation

The final evaluation of the contrast enhancement results has been performed using a full-reference metric. In order to determine which metric is the most suitable for our domain, each of the 14 metrics analyzed in the work of Ponomarenko et al. on the TID2013 dataset [19] has been compared with users Mean Opinion Score only on images distorted under the label "Contrast change". The collected correlation values, based on Spearman and Kendall indices, highlighted the Visual Information Fidelity metric (VIF-P) [20] as the most suitable option for the evaluation of the optimized algorithms performances, with 0.86 and 0.64 correlation values respectively.

Since the used dataset provides for each image multiple enhanced versions but not a target one, it is necessary to select a reference image to perform the comparison. To this end, the following selection procedure has been defined: considering only the positive labels, the image closest to the average contrast/brightness data point has been selected as "average user preferred image", and used as a final reference for evaluation of the optimized contrast enhancement algorithms.

4. EXPERIMENTAL RESULTS

The results for the training of the acceptability regressor are presented in Table 1, following different versions of the data preprocessing procedure. Since the test set presents the same unbalanced nature of the training set in terms of labels, macro accuracy has been used to select the best configuration to be used for the optimization procedure. It is possible to observe how the data balancing procedure contributes drastically to improve macro accuracy, which evaluates both classes equally. The outliers removal procedure introduces a small,

Table 2. Results in terms of VIF-P score and percentage of images from [2] improved after our contrast enhancement.

Optim. level	Contrast enhancement algorithm	VIF-P	% improved images
None	Original input	0.8162	-
None	Local Contrast Correction (LCC)	0.7874	31%
Dataset	Gamma + Histogram Stretch	0.8765	95%
	Gamma + S-curve	0.8166	52%
	Local Contrast Correction (LCC)	0.8645	80%
Image	Gamma + Histogram Stretch	0.8582	85%
	Gamma + S-curve	0.8405	66%
	Local Contrast Correction (LCC)	0.8720	84%

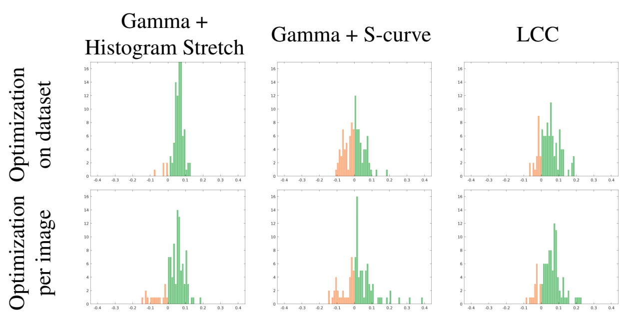

Fig. 3. Distributions of the differences in VIF-P values between the enhanced images and the original input ones.

yet consistent, improvement across different metrics. This suggests that the specific combination of neural features and logistic regressor are inherently robust to label noise.

The final results for optimized contrast enhancement are presented in Table 2 for all three considered algorithms, in terms of average VIF-P before and after enhancement. Two groups of scores are reported, corresponding to the optimization on dataset and per image. As can be seen from the table, the application of the optimization procedure improves the quality of the output images with respect to the input ones. Analyzing in details the three algorithms, different behaviours can be observed. While for both the S-curve and LCC algorithms the per-image optimization brings higher performance with respect to the optimization on dataset, the behaviour with the gamma correction with histogram stretch operation is the opposite. However, with the only exception given by the S-Curve approach optimized on the dataset, the optimization procedure brings an improvement in terms of average VIF-P. The best result in terms of percentage of images whose quality improved after correction is given by the Gamma + Histogram Stretch algorithm, which positively affected 95% of the test images. Finally, the LCC algorithm can be compared against its default parametrization (non-optimized), highlighting the positive impact of its optimization.

A more in detail view of the distribution of differences of quality score between the target images and the ones processed by the three algorithms is shown in Figure 3. As can be seen in this representation, the Gamma + Histogram Stretch algorithm brings the most noticeable improvement, alongside the LCC algorithm in the per-image optimized version.

Fig. 4. Images processed with different optimized contrast enhancement algorithms, to highlight the difference in output enhancements. Other examples available in our repository.

Figure 4 visually presents the effect of modifying input images using the three contrast enhancement algorithms optimized through our framework. It is possible to observe how different algorithms operate differently. Although all the considered algorithms involve non-linear processing, which can introduce modifications in the image chromaticity, the Gamma+S-curve solution appears to be the most strongly affected. These results prove the effectiveness of using user preferences to drive an optimization procedure for contrast and brightness enhancement, and demonstrates how it can also be applied to different kind of algorithms.

5. CONCLUSIONS

We have presented a framework for the optimization of parametric contrast enhancement algorithms. Our solution is based on the definition of an acceptability regression module, trained on user-expressed preferences, which is then exploited as the guiding optimization function for the selection of the best parameters of chosen contrast enhancement algorithms. We have experimented with both dataset-level optimization and image-level optimization, of three different contrast enhancement algorithms of different levels of complexity. Our evaluation shows that the output of the optimized algorithms is consistently improving the quality of the input image, as measured by the Visual Information Fidelity metric. The improvement has been also visually confirmed. In the future, we will collect judgements on the final results by relying to a panel of human observers, and we will expand our analysis to other types of image enhancement [21].

6. REFERENCES

[1] Robert D. Fiete, *Modeling the Imaging Chain of Digital Cameras*, SPIE, Nov. 2010.

[2] Ronnachai Jaroensri, Sylvain Paris, Aaron Hertzmann, Vladimir Bychkovsky, and Fredo Durand, "Predicting range of acceptable photographic tonal adjustments," in *2015 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 2015, pp. 1–9.

[3] Stephen M Pizer, E Philip Amburn, John D Austin, Robert Cromartie, Ari Geselowitz, Trey Greer, Bart ter Haar Romeny, John B Zimmerman, and Karel Zuiderveld, "Adaptive histogram equalization and its variations," *Computer vision, graphics, and image processing*, vol. 39, no. 3, pp. 355–368, 1987.

[4] Yeong-Taeg Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," *IEEE transactions on Consumer Electronics*, vol. 43, no. 1, pp. 1–8, 1997.

[5] Nyamlkhagva Sengee, Altansukh Sengee, and Heung-Kook Choi, "Image contrast enhancement using bi-histogram equalization with neighborhood metrics," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 4, pp. 2727–2734, 2010.

[6] A Mokrane, "A new image contrast enhancement technique based on a contrast discrimination model," *CVGIP: Graphical Models and Image Processing*, vol. 54, no. 2, pp. 171–180, 1992.

[7] Chang-Hsing Lee, Ling-Hwei Chen, and Wei-Kang Wang, "Image contrast enhancement using classified virtual exposure image fusion," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 4, pp. 1253–1261, 2012.

[8] Sebastiano Battiato, Angelo Bosco, Alfio Castorina, and Giuseppe Messina, "Automatic image enhancement by content dependent exposure correction," *EURASIP Journal on Advances in Signal Processing*, vol. 2004, no. 12, pp. 1–12, 2004.

[9] Cheng-Hsiung Hsieh, Bo-Chang Chen, Chih-Ming Lin, and Qiangfu Zhao, "Detail aware contrast enhancement with linear image fusion," in *2010 2nd International Symposium on Aware Computing*. IEEE, 2010, pp. 1–5.

[10] Nathan Moroney, "Local color correction using non-linear masking," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 2000, vol. 2000, pp. 108–111.

[11] Raimondo Schettini, Francesca Gasparini, Silvia Corchs, Fabrizio Marini, Alessandro Capra, and Alfio Castorina, "Contrast image correction method," *Journal of Electronic Imaging*, vol. 19, no. 2, pp. 023005, 2010.

[12] Mahmoud Afifi, Konstantinos G Derpanis, Bjorn Ommer, and Michael S Brown, "Learning multi-scale photo exposure correction," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 9157–9167.

[13] Sing Bing Kang, Ashish Kapoor, and Dani Lischinski, "Personalization of image enhancement," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 2010, pp. 1799–1806.

[14] Gabriel Fillipe Centini Campos, Saulo Martiello Mastelini, Gabriel Jonas Aguiar, Rafael Gomes Mantovani, Leonimer Flávio de Melo, and Sylvio Barbon, "Machine learning hyperparameter selection for contrast limited adaptive histogram equalization," *EURASIP Journal on Image and Video Processing*, vol. 2019, no. 1, pp. 1–18, 2019.

[15] Vladimir Bychkovsky, Sylvain Paris, Eric Chan, and Frédo Durand, "Learning photographic global tonal adjustment with a database of input / output image pairs," in *The Twenty-Fourth IEEE Conference on Computer Vision and Pattern Recognition*, 2011.

[16] Karen Simonyan and Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint*, vol. arXiv:1409.1556, 2014.

[17] Jerome Friedman, Trevor Hastie, and Robert Tibshirani, "Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors)," *The annals of statistics*, vol. 28, no. 2, pp. 337–407, 2000.

[18] Jasper Snoek, Hugo Larochelle, and Ryan P Adams, "Practical bayesian optimization of machine learning algorithms," *Advances in neural information processing systems*, vol. 25, 2012.

[19] Nikolay Ponomarenko, Lina Jin, Oleg Ieremeiev, Vladimir Lukin, Karen Egiazarian, Jaakko Astola, Benoit Vozel, Kacem Chehdi, Marco Carli, Federica Battisti, et al., "Image database tid2013: Peculiarities, results and perspectives," *Signal processing: Image communication*, vol. 30, pp. 57–77, 2015.

[20] Hamid R Sheikh and Alan C Bovik, "Image information and visual quality," *IEEE Transactions on image processing*, vol. 15, no. 2, pp. 430–444, 2006.

[21] Simone Zini and Marco Buzzelli, "On the impact of rain over semantic segmentation of street scenes," in *International Conference on Pattern Recognition*. Springer, 2021, pp. 597–610.