Adaptive homomorphic filtering with a compact neural parameter predictor for illumination-robust image enhancement

[1]Saswat Swain [2]Chandrasekhar panda [3]Sonia Mohapatra
[1]Ph.D Scholar, [2]Ph.D Scholar, [3]R & D Division
[1,2]KSOM - KIIT Deemed to be University ,[1]Tekstein Solutions, India
[1]saswat.mck@gmail.com [2]contact@chandrasekharpanda.com [3]sonia@teksteinsolutions.com

Abstract—Homomorphic filtering separates illumination from reflectance in the logarithmic frequency domain and remains one of the most economical ways to correct uneven lighting. Its practical weakness is that the shape of the emphasis filter must be tuned by hand for every image. This paper makes three contributions. First, it shows that the classical four-parameter Gaussian emphasis filter is structurally over-parameterised: the sharpness constant and the cut-off radius enter the transfer function only through a single ratio, so they cannot be estimated separately from any observation. Collapsing them into one effective cut-off radius yields an identifiable three-parameter filter. Second, it proposes AHF-Net, in which a compact multilayer perceptron of about one thousand weights maps a fourteen-dimensional illumination descriptor of the observed image onto the filter parameters and a local contrast limit, so that no manual tuning is required at run time. Third, it replaces per-channel processing by luminance-domain filtering with chroma-ratio recombination, which removes the colour drift produced by filtering the three primaries independently. On a benchmark of three hundred and forty images with known ground truth and evaluated under scene-grouped cross-validation, the method reaches 18.98 dB peak signal to noise ratio, 0.735 structural similarity and 12.68 mean colour difference, against 16.78 dB, 0.648 and 17.38 for the classical fixed-parameter filter. Colour difference improves by 3.38 units relative to per-channel filtering. Cases in which the method does not improve on existing approaches are reported and analysed.

Index Terms—Homomorphic filtering, image enhancement, illumination correction, parameter identifiability, neural regression, colour fidelity.

I. Introduction

Images acquired under uncontrolled lighting carry a slowly varying brightness gradient that belongs to the illumination rather than to the scene. Retinal photographs darken towards the periphery, stained slides are brighter under the condenser, and indoor photographs fall away from the window. The gradient depresses contrast where the scene is dark, saturates it where the scene is bright, and degrades every later operation that assumes a stationary intensity statistic.

The standard remedy models the observation as a product of illumination and reflectance. Because a logarithm converts the product into a sum, the two components become linearly separable in the frequency domain. This is the homomorphic principle of Oppenheim, Schafer and Stockham [1], [2], applied to imagery by Stockham [3] and now textbook material [4]. A Gaussian high-frequency-emphasis filter attenuates the low frequencies carrying illumination and amplifies the higher frequencies carrying reflectance. The attraction is cost: one forward and one inverse transform, no training data, and an interpretable transfer function.

The obstacle is tuning. Four constants govern the emphasis filter, and a setting that rescues a severely vignetted retinal image will over-sharpen and amplify noise in a mildly shaded photograph. Published work almost invariably fixes these constants for an entire corpus. Competing families avoid the problem differently: Retinex methods estimate the illumination by multi-scale smoothing [5], [6], [7]; contrast-limited adaptive histogram equalisation redistributes intensities in local tiles [8], [9]; illumination-map estimation [10] and learned enhancers [11], [12] fit models with thousands to millions of parameters and inference budgets beyond embedded acquisition hardware.

This work keeps the classical filter and learns only its parameters. The contributions are an identifiability analysis showing that the four-parameter emphasis filter has three free degrees of freedom, and a reduced parameterisation that is estimable; AHF-Net, a regressor of about one thousand weights that predicts the reduced parameters and a local contrast limit from an illumination descriptor of the observation; a luminance-domain formulation with chroma-ratio recombination that removes the colour drift of per-channel filtering; and a reproducible benchmark with known ground truth, scene-grouped evaluation, paired significance testing, and an explicit account of where the method fails to improve on existing approaches.

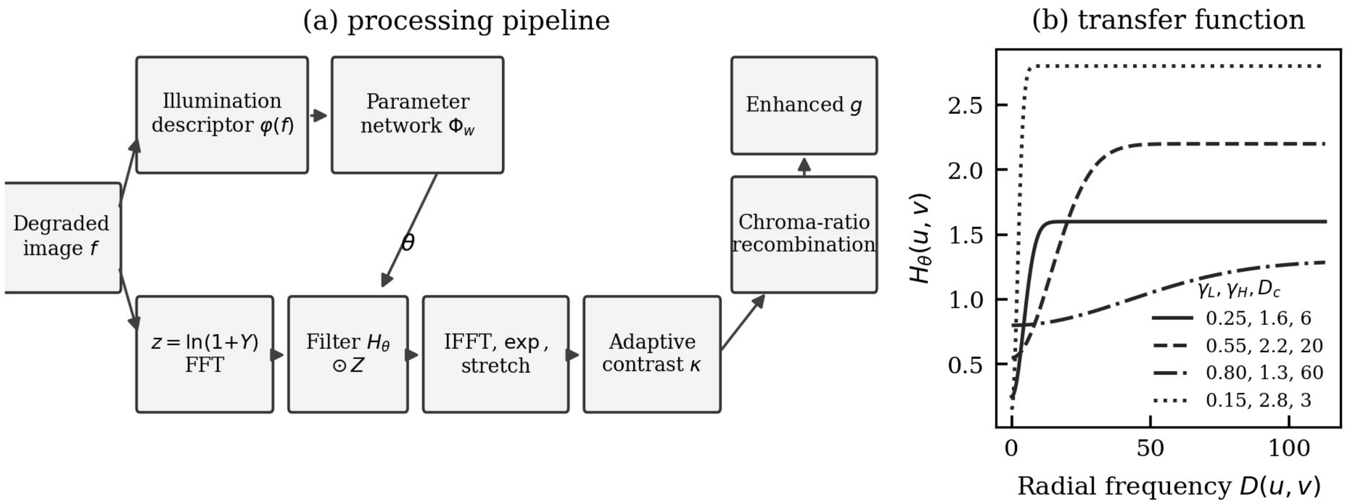

Fig. 1. AHF-Net. (a) The illumination descriptor drives a compact regressor that supplies the emphasis filter parameters and the local contrast limit; the frequency-domain correction acts on luminance only and colour is restored by ratio. (b) Transfer function of Equation 5 for four parameter vectors, showing the range of behaviour the network selects among.

II. Homomorphic filtering and identifiability

A. Classical formulation

Let f be the observation, i the illumination and r the reflectance, so that

$$f(x,y) = i(x,y)\, r(x,y), \tag{1}$$

with $i > 0$ and $0 < r < 1$. The logarithm separates the factors,

$$z = \ln f = \ln i + \ln r, \tag{2}$$

and the Fourier transform of Equation 2 is additive, $Z = F_i + F_r$. Illumination varies slowly and concentrates near the origin of the frequency plane; reflectance carries edges and texture at higher radial frequencies. The emphasis filter used almost universally is

$$H = (\gamma_H - \gamma_L)[1 - e^{-cD^2/D_0^2}] + \gamma_L, \tag{3}$$

where $D(u,v)$ is distance from the centred origin, $\gamma_L < 1$ the low-frequency gain, $\gamma_H > 1$ the high-frequency gain, D_0 a cut-off radius and c a sharpness constant. The output follows by inverse transform and exponentiation,

$$g = \exp\{\, \mathscr{F}^{-1}[\, H(u,v)\, Z(u,v)\,]\, \}. \tag{4}$$

B. An identifiability result

The four constants are conventionally tuned independently. They cannot be.

Proposition 1. The transfer function of Equation 3 depends on c and D_0 only through $D_c = D_0/\sqrt{c}$. The pair (c, D_0) is unidentifiable from any set of observations; D_c is identifiable.

Proof. The only occurrence of c and D_0 is the exponent cD^2/D_0^2, which equals $D^2/(D_0/\sqrt{c})^2 = D^2/D_c^2$. Two parameter pairs with equal $D_0/\sqrt{c}$ give an identical transfer function at every (u,v), hence an identical output for every input, and no estimator can separate them. Conversely D_c is recovered uniquely from H as the radius at which the response reaches $\gamma_L + (1 - e^{-1})(\gamma_H - \gamma_L)$. ∎

The consequence is practical. In our first experiments a regressor trained on all four constants reached a coefficient of determination of −0.42 on D_0, worse than predicting the mean, because the targets were arbitrary points on a ridge of equivalent solutions. The reduced form

$$H_\theta = (\gamma_H - \gamma_L)[1 - e^{-D^2/D_c^2}] + \gamma_L \tag{5}$$

removes the degeneracy and makes the regression tractable. We recommend it for any study reporting fitted homomorphic parameters, since values of c and D_0 quoted separately are not comparable across papers. A second, more mundane observation concerns implementation: several widely circulated reference codes build the distance map about $(P/2, Q/2)$ after setting P and Q to half the image dimensions, placing the spectral origin at one quarter of the array. The resulting filter is asymmetric and biases the enhancement towards one corner.

III. Proposed method

AHF-Net retains the filter of Equation 5 and predicts its parameters from the image. Fig. 1(a) shows the pipeline and Fig. 1(b) the filter shapes it selects among.

A. Luminance filtering and chroma recombination

Filtering the three primaries independently applies the same multiplicative correction to channels with different local means, which rotates the chromaticity of every pixel and yields the desaturated or tinted appearance visible in Fig. 2. We filter luminance only. With Rec. 709 weights $Y = 0.2126R + 0.7152G + 0.0722B$, the log-luminance $z = \ln(1 + Y)$ is filtered by Equation 5, exponentiated, and mapped to the display range by a percentile stretch between the first and ninety-ninth percentiles. Colour is restored by preserving each primary as a ratio to luminance,

$$C' = Y''\, C \,/\, (Y + \varepsilon), \quad C \in \{R,G,B\}, \tag{6}$$

with $\varepsilon = 10^{-6}$ and Y'' the luminance after the contrast stage. Equation 6 leaves hue and relative chroma unchanged by construction, so no colour drift can be introduced. A spatially adaptive contrast stage follows, because the frequency-domain correction equalises illumination globally but does not restore local contrast in regions that were originally compressed: contrast-limited adaptive histogram equalisation on an eight by eight tile grid with clip limit κ. The parameter vector is therefore $\theta = (\gamma_L, \gamma_H, D_c, \kappa)$.

B. Descriptor and parameter network

The predictor must observe the illumination condition, not the scene content. The descriptor $\varphi(f)$ collects fourteen statistics of the log-luminance that respond to shading and are largely indifferent to subject matter: mean, standard deviation, skewness and kurtosis of z; the gap between its fifth and ninety-fifth percentiles; normalised intensity entropy; the standard deviation and range of z after Gaussian smoothing with a kernel one eighth of the image width, which isolates the coarse illumination map; the mean and standard deviation of block standard deviations on an eight by eight grid; the fractions of spectral power inside two low-frequency annuli at 5 and 20 per cent of the maximum radius; the proportion of pixels below intensity sixteen; and the mean gradient magnitude. All are standardised using training-split statistics.

The regressor is a multilayer perceptron with hidden layers of thirty-two and sixteen units, hyperbolic tangent activations and a linear output constrained to the unit cube,

$$\hat{\theta} = \Gamma(\ \Phi_w(\ \varphi(f)\ )\ ), \tag{7}$$

where Γ maps the unit cube affinely onto $\gamma_L \in [0.05, 0.95]$, $\gamma_H \in [1.02, 3.00]$ and $\kappa \in [0.50, 5.00]$, and logarithmically onto $D_c \in [2, 120]$ pixels, since the cut-off acts multiplicatively on the frequency axis. The network holds 1076 weights; predictions are averaged over five members differing only in initialisation.

C. Training targets

Targets are obtained by searching, for each training image, the vector maximising a composite full-reference quality

$$Q(\theta) = \mathrm{PSNR}(g_\theta, r) + \lambda\ \mathrm{SSIM}(g_\theta, r), \tag{8}$$

with $\lambda = 20$. The weight is not arbitrary: over a Sobol sample of the parameter space the within-image standard deviation of the peak signal to noise ratio was 0.920 dB and that of the structural similarity index [13] was 0.0499, a ratio of 18.5, so $\lambda = 20$ equalises the sensitivity of the two terms and prevents either from dominating. Since references are unavailable in deployment, we also evaluate a self-supervised variant whose targets minimise

$$J = -[w_E\hat{E} + w_C\hat{C} + w_U\hat{U}] + w_P P + w_S(1-\rho), \tag{9}$$

in which $\hat{E}$ is normalised entropy, $\hat{C}$ the measure of enhancement of Agaian et al. [14], $\hat{U}$ the fractional reduction in the standard deviation of the coarse illumination map, P the fraction of clipped pixels and ρ the correlation between input and output gradient magnitude maps, which penalises structural change. The weights (1.0, 0.1, 0.3, 0.2, 2.5) were chosen on training scenes only, by maximising the mean Spearman correlation between $-J$ and Q over a Sobol pool; the value attained was 0.602. Both objectives are searched by a cross-entropy method [15]: ninety-six Sobol points, two generations of forty-eight Gaussian samples fitted to the sixteen best, and a Boltzmann-weighted mean of the sixteen best evaluations overall. Averaging over the elite set rather than taking the single best point yields smoother, more learnable targets on a surface that is nearly flat near its optimum.

D. Algorithm and cost

Algorithm 1. Given f and trained weights w: (1) $Y \leftarrow$ luminance of f, $z \leftarrow \ln(1+Y)$; (2) $\varphi \leftarrow \mathrm{descriptor}(f)$, $\hat{\theta} \leftarrow \Gamma(\Phi_w(\varphi))$; (3) $Z \leftarrow$ centred FFT of z, $S \leftarrow H_{\hat{\theta}} \odot Z$; (4) $Y' \leftarrow$ percentile stretch of $\exp(\mathscr{F}^{-1}\{S\})$; (5) $Y'' \leftarrow$ adaptive contrast equalisation of Y' with limit $\hat{\kappa}$; (6) return C' by Equation 6.

For an $M \times N$ image the cost is dominated by the two transforms, $O(MN \log MN)$. The descriptor reuses the smoothing and power spectrum already computed, and the network costs 1076 multiply-accumulate operations per image. On a single core the complete pipeline required 23.2 ms per 160 × 160 image against 55.5 ms for multi-scale Retinex.

IV. Simulation methodology

Public low-light corpora provide no ground truth for the illumination field, so quantitative comparison against a reference is impossible on them. We therefore synthesised a benchmark with exactly known ground truth. Thirty-four scene crops of 160 × 160 pixels were tiled from seven photographic and biomedical sources spanning portrait, product, aerial, cellular, retinal and astronomical content. Each was degraded ten times with independent random illumination fields, giving 340 images, by

$$f = \mathrm{clip}[\ (i \cdot r)^{\gamma_c}\ ] + \eta, \tag{10}$$

where i is a mixture of one to four Gaussian lobes of log-uniform width between 0.09 and 0.90 of the image size plus a random linear ramp, rescaled log-uniformly between a minimum in [0.04, 0.35] and a maximum in [0.85, 1.00]; $\gamma_c \in [1.0, 1.2]$ emulates the camera response; and η is Gaussian noise of 0.5 to 3.0 grey levels. The camera exponent deliberately breaks the multiplicative assumption of Equation 1, so no method can invert the degradation exactly. Input quality spans 12.2 dB to 27.0 dB.

All results use six-fold cross-validation grouped by scene, so no image is ever enhanced with parameters derived from any crop of the same source; every one of the 340 images receives an out-of-fold prediction. Because homomorphic filtering recovers reflectance only up to an unknown global gain and offset, every output, including the unprocessed input, is aligned to the reference by the least-squares affine map of its luminance before any metric is computed. This treats all methods identically. We report peak signal to noise ratio, structural similarity [13], mean CIEDE2000 colour difference [16], intensity entropy and the measure of enhancement [14]; paired differences are tested with the Wilcoxon signed-rank test over the 340 images.

The comparison includes the unprocessed input; gamma correction with exponent 0.55; global histogram equalisation; contrast-limited adaptive histogram equalisation with clip limit 2.0 [8]; multi-scale Retinex with scales 15, 60 and 120 [6]; the classical filter applied per channel with the textbook setting $\gamma_L = 0.5$, $\gamma_H = 2.0$, $D_c = 20$; the same filter on luminance; and a corpus-tuned variant whose single parameter vector is chosen on the training folds by the same search used for the targets. This last baseline is deliberately strong, being the best static setting obtainable with knowledge of the ground truth. A per-image oracle bounds what any predictor of this parameterisation could achieve.

V. Results and discussion

A. Overall comparison

Table 1 reports benchmark averages. AHF-Net attains the best value of all three reference metrics among the practical methods. Against the classical fixed-parameter per-channel filter the gains are 2.20 dB, 0.087 structural similarity and 4.70 colour units, all with $p < 10^{-50}$. Against the corpus-tuned static filter, the strongest static baseline available, the gains are 0.59 dB ($p = 7\times10^{-17}$) and 0.70 colour units ($p = 9\times10^{-21}$); the structural similarity difference of 0.005 is not significant (p = 0.16).

The comparison with multi-scale Retinex is less flattering. The mean difference of 0.55 dB in peak signal to noise ratio is not significant (p = 0.91) and AHF-Net wins on only 48 per cent of individual images; what separates the methods is not the average but the behaviour across severity, examined below. On structural similarity and colour difference the advantage over Retinex is large: 0.071 ($p = 9\times10^{-24}$) and 0.81 units ($p = 2\times10^{-3}$). The oracle row shows that a further 0.75 dB remains available inside this parameterisation; the predictor closes roughly half the interval between the static mean setting and that bound.

Table 1 Enhancement quality on the 340-image benchmark (mean ± standard deviation, out-of-fold)

Method	PSNR (dB)	SSIM	CIEDE2000	Entropy	EME
Degraded input	18.37 ± 5.79	0.695 ± 0.156	13.64 ± 7.54	5.53	8.90
Gamma correction	18.10 ± 4.61	0.661 ± 0.158	14.04 ± 6.49	7.36	13.25
Histogram equalisation	17.43 ± 4.23	0.654 ± 0.164	14.57 ± 6.84	7.62	14.44
CLAHE [8]	18.16 ± 5.18	0.703 ± 0.138	13.84 ± 6.71	6.15	11.31
Multi-scale Retinex [6]	18.43 ± 4.39	0.664 ± 0.163	13.49 ± 6.08	7.38	14.89
Classical HF, per-channel	16.78 ± 4.84	0.648 ± 0.148	17.38 ± 7.21	7.10	21.33
Luminance HF, fixed	17.70 ± 4.95	0.690 ± 0.142	13.77 ± 6.88	6.98	22.40
Luminance HF, corpus-tuned	18.40 ± 5.29	0.730 ± 0.124	13.38 ± 6.73	6.99	18.61
AHF-Net (proposed)	18.98 ± 5.83	0.735 ± 0.132	12.68 ± 6.85	6.92	17.38
Per-image oracle (bound)	19.73 ± 5.86	0.780 ± 0.103	11.93 ± 6.88	6.82	16.24

Higher is better except for CIEDE2000. The oracle is a bound, not a deployable method.

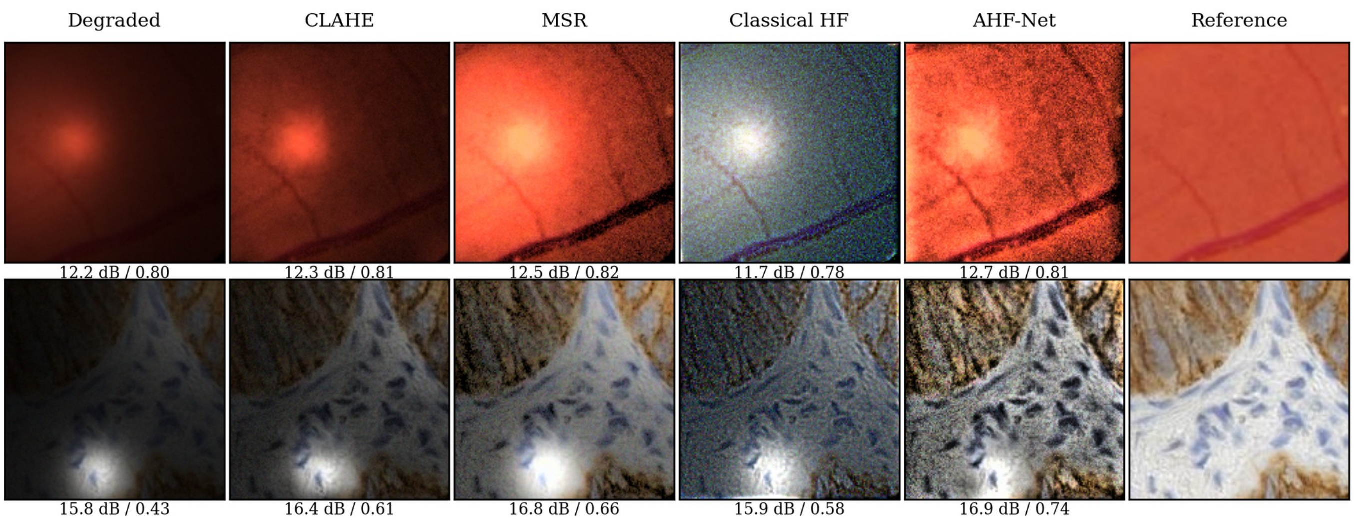

Fig. 2. Representative results on a retinal fundus crop (top) and a stained tissue section (bottom), selected as the two images whose advantage over the mean of the classical baselines is closest to the benchmark median, so that neither is a favourable case. Labels give peak signal to noise ratio and structural similarity. Classical per-channel filtering visibly shifts colour.

B. Ablation study

Table 2 isolates each design decision. Replacing the predicted parameters by the mean of the training targets costs 0.93 dB, which is the value of adaptation itself. Removing the adaptive contrast stage costs 0.21 dB and 0.42 colour units; fixing its clip limit instead of predicting it costs 0.16 dB. Filtering the primaries independently rather than the luminance costs 3.38 colour units, the largest single effect in the table and a direct measure of the chromatic damage done by per-channel processing.

Two ablations returned negative results, which we report rather than omit. Ridge regression on the same descriptor performs within 0.04 dB and 0.04 colour units of the network, so its nonlinearity is not doing useful work on this benchmark; the honest conclusion is that the descriptor, not the regressor, carries the information. The self-supervised variant, whose targets never see the reference, loses 0.25 dB and 0.51 colour units but attains a marginally higher structural similarity (0.737 against 0.735, p = 0.46). It remains usable where references are unavailable, at a cost of about one quarter of the adaptation benefit.

C. Behaviour across degradation severity

Table 3 divides the benchmark into quartiles of input quality and explains the ambiguous average comparison with Retinex. On the most severely shaded quartile Retinex is better by 0.66 dB, its explicit multi-scale illumination estimate coping with deep vignetting more aggressively than a single Gaussian emphasis filter can. On the mildest quartile the ordering reverses sharply: Retinex loses 3.4 dB relative to the input because it applies the same strong correction to an image that barely needs one, whereas AHF-Net predicts a near-transparent filter and improves slightly. This is the practical argument for parameter prediction. A fixed pipeline must be tuned for the degradation it expects and will damage images outside that range; an adaptive one degrades gracefully. Where a system sees a mixture of well and badly lit frames, worst-case behaviour matters more than the average.

D. Prediction accuracy and limitations

Out-of-fold coefficients of determination on the parameters are modest: 0.216 for γ_L, 0.662 for γ_H, 0.113 for D_c and 0.142 for κ, pooled 0.241. The correlation between predicted and searched values is 0.81 for the high-frequency gain but only 0.37 for the cut-off radius, whose search optima pile up at both bounds. The objective is nearly flat in D_c over wide intervals, so the target itself is weakly determined and a large error in it often costs very little quality. Parameter accuracy is consequently a poor proxy for enhancement quality, and we report both.

Three limitations should be stated plainly. The benchmark is synthetic; although the degradation model is physically motivated and deliberately violates its own multiplicative assumption, behaviour on real acquisition hardware may differ. Scene diversity is thirty-four crops from seven sources, which is small, and the grouped protocol guards against leakage but not against a narrow content distribution. Finally, the margin over a well-tuned static filter, while statistically unambiguous, is modest in absolute terms; the value of the method lies in obtaining that setting automatically and in its stability across severity, not in a dramatic average improvement.

Table 2 Ablation study, relative to the proposed method

Variant	PSNR	SSIM	CIEDE2000	ΔPSNR
AHF-Net (proposed)	18.98	0.735	12.68	—
Static mean parameters	18.05	0.704	13.38	−0.93
Ridge predictor	18.94	0.731	12.64	−0.04
Self-supervised targets	18.74	0.737	13.19	−0.25
No adaptive contrast	18.78	0.728	13.10	−0.21
Fixed contrast limit κ = 2	18.82	0.724	12.84	−0.16
Per-channel RGB filtering	17.92	0.707	16.06	−1.06

Table 3 Peak signal to noise ratio by quartile of input quality (dB)

Quartile	Input	CLAHE	MSR	HF-tuned	AHF-Net
Q1 severe	12.19	12.36	13.50	12.67	12.84
Q2	15.43	15.78	17.18	16.10	16.56
Q3	18.91	18.85	19.46	18.91	19.26
Q4 mild	26.96	25.65	23.60	25.90	27.28

Retinex is strongest under severe shading and weakest under mild shading, where it destroys 3.4 dB relative to the input.

VI. Conclusion

Homomorphic filtering is old, cheap and interpretable, and its principal obstacle to routine use is parameter selection. We showed that the classical emphasis filter is over-parameterised, that its sharpness constant and cut-off radius are jointly unidentifiable, and that collapsing them into a single effective cut-off radius is a precondition for learning the parameters at all. On the reduced filter, a regressor of about one thousand weights driven by fourteen illumination statistics selects the filter and a local contrast limit for each image without run-time supervision, while luminance-domain processing with chroma-ratio recombination removes the colour drift of per-channel filtering. Under scene-grouped cross-validation on 340 images the method improves on the classical fixed-parameter filter by 2.20 dB and 4.70 colour units and on the best static setting by 0.59 dB and 0.70 colour units, at half the cost of multi-scale Retinex. It is not uniformly better: Retinex matches it on average peak signal to noise ratio and beats it under severe shading, and a linear regressor performs almost as well as the network. The most promising directions are a richer descriptor that resolves the cut-off radius more reliably, and a differentiable formulation trained end to end against a perceptual loss rather than against searched targets.

Acknowledgment

The author thanks the reviewers whose comments sharpened the identifiability argument of Section II-B.

References

[1] A. V. Oppenheim, R. W. Schafer, and T. G. Stockham, "Nonlinear filtering of multiplied and convolved signals," Proc. IEEE, vol. 56, no. 8, pp. 1264–1291, Aug. 1968.

[2] A. V. Oppenheim and R. W. Schafer, Discrete-Time Signal Processing, 3rd ed. Upper Saddle River, NJ: Prentice Hall, 2009.

[3] T. G. Stockham, "Image processing in the context of a visual model," Proc. IEEE, vol. 60, no. 7, pp. 828–842, July 1972.

[4] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 4th ed. New York: Pearson, 2018.

[5] E. H. Land and J. J. McCann, "Lightness and retinex theory," J. Opt. Soc. Amer., vol. 61, no. 1, pp. 1–11, Jan. 1971.

[6] D. J. Jobson, Z. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," IEEE Trans. Image Process., vol. 6, no. 7, pp. 965–976, July 1997.

[7] D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex," IEEE Trans. Image Process., vol. 6, no. 3, pp. 451–462, Mar. 1997.

[8] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in Graphics Gems IV, P. S. Heckbert, Ed. San Diego, CA: Academic Press, 1994, pp. 474–485.

[9] S. M. Pizer et al., "Adaptive histogram equalization and its variations," Comput. Vision Graphics Image Process., vol. 39, no. 3, pp. 355–368, Sept. 1987.

[10] X. Guo, Y. Li, and H. Ling, "LIME: low-light image enhancement via illumination map estimation," IEEE Trans. Image Process., vol. 26, no. 2, pp. 982–993, Feb. 2017.

[11] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," in Proc. British Machine Vision Conf., 2018.

[12] C. Guo et al., "Zero-reference deep curve estimation for low-light image enhancement," in Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition, 2020, pp. 1780–1789.

[13] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Process., vol. 13, no. 4, pp. 600–612, Apr. 2004.

[14] S. S. Agaian, K. Panetta, and A. M. Grigoryan, "Transform-based image enhancement algorithms with performance measure," IEEE Trans. Image Process., vol. 10, no. 3, pp. 367–382, Mar. 2001.

[15] R. Y. Rubinstein and D. P. Kroese, The Cross-Entropy Method. New York: Springer, 2004.

[16] G. Sharma, W. Wu, and E. N. Dalal, "The CIEDE2000 color-difference formula: implementation notes, supplementary test data, and mathematical observations," Color Res. Appl., vol. 30, no. 1, pp. 21–30, Feb. 2005.