Convergence of the Variational Approach to Exact Histogram Specification for N-D Images

Jeffrey Bush - bushj@moravian.edu

November 2022

The variation approach (VA) method [1, 4, 3] for exact histogram specification / equalization is only defined for 2D images. This is a powerful method for exact histogram specification but it would be great if it could work for arbitrary N-dimensional images as well.

To extend the VA method into higher dimensions, the neighborhood must include additional pixels, the gradient matrix must use these pixels and the parameters β, α_1, and α_2 must be adjusted so that all necessary convergence constraints are still met. The following constraints are defined across the various papers:

$$\beta \in (0, {}^1\!/\!_\eta) \qquad [1, (20)] \tag{1}$$

$$\theta'\left(v_f - 2\xi(\beta\eta, \alpha_1), \alpha_2\right) \in (1-\epsilon, 1), \epsilon \approx 10^{-5} \qquad [4, (14)] \tag{2}$$

$$8\beta\xi'(4\beta, \alpha_1)\theta''(0, \alpha_2) < 1 \qquad [3, (10)] \tag{3}$$

Using $\theta(t,\alpha) = |t| - \alpha\log(1 + {}^{|t|}\!/\!_\alpha)$ chosen in [3], we have $\theta'(t,\alpha) = {}^t\!/\!_{\alpha+|t|}$ and $\xi = (\theta')^{-1} = {}^{\alpha t}\!/\!_{1-|t|}$. The value v_f is the maximal minimal neighbor difference in the input image as defined in [1, (14)] and η is the sum of all the neighbor weights γ_i as defined in [1, (10)]. The values of γ_i (and thus η) have a large influence on the v_f value as shown in [4, 1] and thus influence the second constraint.

When the third constraint was added in [3] for the new minimization technique, the value of η was fixed to 4 and removed from equations along with the neighbor weights γ_i. Thus, to be able to use this constraint with a different number of neighbors or weights, as would be necessary for higher-dimensional images, it has to be re-derived. This requires an expansion of the G matrix definition originally defined in [3] as a stack of sub-matrices G_j with each sub-matrix representing the differences of each pixel with the neighbors in a pair of directions (e.g. left and right). The original only describes the sub-matrices for the horizontal and vertical directions. This can be extended by adding additional sub-matrices for the diagonals and/or including the depth direction in 3D images. Each of these sub-matrices can also be multiplied by $\hat{\gamma_j}$ so that each direction has a different weighting. Since each sub-matrix is a pair of directions, there is a single $\hat{\gamma_j}$ for each, which matches a pair of γ_i values. The G matrix for a $P \times N \times M$ image with a minimal 6-neighborhood taken column-wise would be:

$$G = \begin{pmatrix} \hat{\gamma_h} G_h \\ \hat{\gamma_v} G_v \\ \hat{\gamma_d} G_d \end{pmatrix} = \begin{pmatrix} \hat{\gamma_h} I_M \otimes I_N \otimes D_P \\ \hat{\gamma_v} I_M \otimes D_N \otimes I_P \\ \hat{\gamma_d} D_M \otimes I_N \otimes I_P \end{pmatrix} \in \mathbb{R}^{r,n} \tag{4}$$

where I_N is the $N \times N$ identity matrix, D_N is the $(N-1) \times N$ forward difference matrix [4], $\otimes$ denotes the Kronecker product, $n = MNP$ is the total number of voxels in the image, $r = 3MNP - MN - MP - NP$, and $\hat{\gamma_h}, \hat{\gamma_v}, \hat{\gamma_d}$ are the neighbor weights for the horizontal, vertical, and depth directions. In this setup, $\eta = 2 \sum_j \hat{\gamma_j} = \sum_i \gamma_i$. This concept can be extended to diagonal directions and higher dimensions.

For any of the sub-matrices $||G_j^T G_j||_2 \leq 4$ since the $G_j^T G_j$ matrix is a Lapacian matrix with the values along the diagonal being primarily 2 (some values are 0 and 1 for corner and edge pixels) since every pixel is used twice: once with a $+1$ and once with a -1. [2] showed that the upper bound for the largest eigenvalue (and thus the ℓ_2-norm) of a Lapacian matrix is the maximum of the sum of two adjacent vertices and thus $||G_j^T G_j||_2 = ||G_j||_2^2 \leq 2 + 2 = 4$.

The upper bound of the ℓ_2-norm of G can then be derived using the definition of ℓ_2-norm: $||A||_2 = \sup_{||x||_2=1} ||Ax||_2$.

$$||G||_2^2 = \sup_{||x||_2=1} ||Gx||_2^2 = \sup_{||x||_2=1} x^T G^T G x \tag{5}$$

$$= \sup_{||x||_2=1} \sum_j \hat{\gamma_j}^2 x^T G_j^T G_j x \tag{6}$$

$$\leq \sum_j \hat{\gamma_j}^2 \sup_{||x||_2=1} x^T G_j^T G_j x = \sum_j \hat{\gamma_j}^2 ||G_j||_2^2 \tag{7}$$

And thus:

$$||G||_2^2 \leq \sum_j \hat{\gamma_j}^2 ||G_j||_2^2 \leq 4 \sum_j \hat{\gamma_j}^2 = 2 \sum_i \gamma_i^2 \tag{8}$$

For simplicity, and due to its similarity to η, $\bar{\eta} := \sum_i \gamma_i^2$ and thus $||G||_2^2 = 2\bar{\eta}$.

Similarly we can compute the upper bound of the ℓ_1- and ℓ_∞-norms: $||G||_1 \leq 2 \sum_j \hat{\gamma_j} = \eta$ and $||G||_\infty \leq 2 \max_j \hat{\gamma_j}$. These can be seen intuitively since the vast majority of columns in the G_j sub-matrices have a single $+1$ and -1 multiplied by the respective $\hat{\gamma_j}$ and each row has a single $+1$ and -1 multiplied by the respective $\hat{\gamma_j}$.

Using these we can turn back to the third constraint and derive it using the same process as in [3, Theorem 3] but with the more detailed values for $||G||_1$, $||G||_2$, and $||G||_\infty$. This results in the following condition to update [3, (10)]:

$$2\bar{\eta}\beta\xi'(\eta\beta, \alpha_1)\,\theta''(0, \alpha_2) < 1 \tag{9}$$

Substituting the fixed values for $\theta(t, \alpha)$ and $\xi(t, \alpha)$ it becomes:

$$\frac{\alpha_1}{\alpha_2} \frac{2\bar{\eta}\beta}{(1 - \eta\beta)^2} < 1 \tag{10}$$

In [3, Remark 1] they state this constraint is overly restrictive due to the assumptions of the image being infinitely large with every pixel having the same value. For the values of $\beta = 0.1$ and $\alpha_1 = \alpha_2 = 0.5$ chosen in [3] for $\eta = 4$, the third constraint comes to $2.2\bar{2}$ which was considered satisfactory to achieve convergence even though not less than 1.

When moving to 3D the values need to be adjusted. If each pixel is equally connected to its immediate neighbors, $\eta = 6$. The first constraint is still satisfied with $\beta = 0.1$ since $\beta \in (0, 1/\eta) = (0, 0.1\bar{6})$ but the third constraint becomes significantly worse, coming in at 7.5. Noticing the connection between η and β indicates that β should be scaled based on η. If linearly scaling, this results in a value of $\beta = \frac{4}{6}0.1 = 0.0\bar{6}$. Scaling β, however, causes the second constraint to move further from the target value of 1. The values of α_1 and α_2 can also be linearly scaled using η resulting in $\alpha_1 = \alpha_2 = \frac{4}{6}0.05 = 0.0\bar{3}$. With these values, all 3 constraints are satisfied as well as they are in [3].

Since each direction can be differently weighted means that anisotropic data can be handled appropriately by changing the weights based on the distances between the pixels. For example, if each plane of a 3D image was further away than pixels within a single plane the $\hat{\gamma}_v$ could be lowered reducing the importance of the depth dimension. The linear scaling of parameters approach used above would likely not work as well when all weights are not 1 and the tuning of parameters will have to be done more carefully to ensure convergence.

Acknowledgments

The author would like to express his gratitude to his colleague Dr. Shannon Talbot for her feedback and discussions on the convergence criteria of the VA method which helped determine the necessary constants for multi-dimensional images.

References

[1] F. Baus, M. Nikolova, and G. Steidl. Fully smoothed $\ell_1 - TV$ models: Bounds for the minimizers and parameter choice. *J. of Mathematical Imag. and Vision*, 48(2):295–307, Feb 2013.

[2] William N. Anderson Jr. and Thomas D. Morley. Eigenvalues of the laplacian of a graph. *Linear and Multilinear Algebra*, 18(2):141–145, 1985.

[3] M. Nikolova and G. Steidl. Fast ordering algorithm for exact histogram specification. *IEEE Trans. Image Process.*, 23(12):5274–5283, Dec 2014.

[4] Mila Nikolova, You-Wei Wen, and Raymond Chan. Exact histogram specification for digital images using a variational approach. *J. of Mathematical Imag. and Vision*, 46(3):309–325, Jul 2013.