Equivalence of histogram equalization, histogram matching and the Nyul algorithm for intensity standardization in MRI

Jesse Knight — University of Guelph, ON, Canada
Graham Taylor — University of Guelph, ON, Canada and Vector Institute
April Khademi — Ryerson University, ON, Canada

Abstract

Intensity standardization is an important preprocessing step in automated analysis of MRI. A popular method by Nyul et al. uses a piece-wise linear approximation of histogram matching. We show that this method is a non-uniform trapezoidal Riemann approximation of the ideal histogram matching operation, and suggest that histogram matching is no better than histogram equalization for intensity standardization in MRI. Experimental results are derived using synthetic data.

1 Introduction

One major obstacle to automated analysis of magnetic resonance images (MRI) is the lack of standardization of the intensity scale [1]. Tissue intensities vary with the acquiring scanner (model, field strength) and image weighting (selection of TE, TR, etc.). As a result, many tools for analyzing images from multiple scanners employ some method of intensity standardization as preprocessing.

For an image $\mathcal{X}$ with intensities denoted x, an intensity transformation can be generally defined as $\tau : x \mapsto y$, $y = \tau(x)$, yielding the adjusted image $\mathcal{Y}$. In the context of standardization, τ is usually monotonic, and uses characteristics of each realization of $\mathcal{X}$ to minimize variation across a dataset. Many methods for intensity standardization have been proposed. The work by Nyul et al. [1, 2] describes one such method, which uses a piece-wise approximation of histogram matching.

In this work, we will show that the transformation proposed by Nyul et al. is specifically a non-uniform trapezoidal Riemann approximation of the ideal histogram matching operation. We will also show that errors associated with this approximation introduce artifacts in the matched histograms, motivating the use of the ideal histogram matching operation as a superior alternative. Experiments using simulated data illustrate the convergence of the approximation for large N, and the superiority of ideal histogram matching.

2 Histogram Equalization

Histogram equalization is one classic method of intensity standardization. The histogram of an image $\mathcal{X}$ comprising K pixels with intensities $x \in [x_{\min}, x_{\max}]$ is denoted h_x; it represents the number of occurrences of each intensity in the image. The intensity probability mass function (PMF) $p_x(x)$ is therefore given by h_x/K,

$$p_x(x) = \frac{1}{K} h_x(x) = \frac{1}{K} \sum_{k=1}^{K} \begin{cases} 1 & \mathcal{X}(k) = x \\ 0 & \mathcal{X}(k) \neq x \end{cases} \tag{1}$$

The cumulative distribution function (CDF) $P_x(x)$ is simply the cumulative sum of $p_x(x)$,

$$P_x(x) = \sum_{x_{\min}}^{x} p_x(\delta) d\delta. \tag{2}$$

Now, consider the intensity transformation given by the CDF of $\mathcal{X}$,

$$\begin{aligned} y &= \tau_x(x) \\ &= P_x(x). \end{aligned} \tag{3}$$

The result of this transformation is $\mathcal{Y}$, and its PMF is denoted p_y. From probability theory, the PMF of a transformed variable is defined as

$$p_y(y) = p_x(x) \left| \frac{dx}{dy} \right|. \tag{4}$$

Using calculus, we can see that

$$\begin{aligned} \frac{dy}{dx} &= \frac{d}{dx} P_x(x) \\ &= \frac{d}{dx} \left[\sum_{x_{\min}}^{x} p_x(\delta) d\delta \right] \\ &= p_x(x), \end{aligned} \tag{5}$$

which means that the PMF of $\mathcal{Y}$,

$$\begin{aligned} p_y(y) &= p_x(x) \left| \frac{dx}{dy} \right| \\ &= p_x(x) \left| \frac{1}{p_x(x)} \right| \\ &= 1, \quad x \in [x_{\min}, x_{\max}] \end{aligned} \tag{6}$$

is a uniform distribution. Thus, using the intensity CDF of $\mathcal{X}$ as an intensity transformation is a *histogram equalization* operation.

3 Histogram Matching

Histogram matching is another method of standardization. A histogram matching transformation $\tau_{xz} : x \mapsto z$ takes any input image $\mathcal{X}$ and yields an output image $\mathcal{Z}$ with PMF p_z. Following on the results above, this operation can be defined as the function composition of a histogram equalization transform $\tau_x = P_x$ and the inverse equalization transform $\tau_z^{-1} = P_z^{-1}$,

$$\tau_{xz} = P_z^{-1}(P_x) \tag{7}$$

Considering this result, we observe that P_z^{-1} does not depend on $\mathcal{X}$ – i.e. p_z is independent of p_x – and is therefore applied equally to all images after the image-specific equalization τ_x is applied. Therefore, the target PMF p_z is unrelated to the objective of intensity standardization across different realizations of $\mathcal{X}$. That is, histogram matching equates to histogram equalization in the task of intensity standardization.

4 Histogram Equalization Approximation

We now turn to the intensity transformation presented in [1]. We show how this transformation is a non-uniform trapezoidal Riemann approximation of the ideal histogram matching operation.

The proposed method uses histogram "landmarks" to define a piece-wise linear intensity transformation for each image $\mathcal{X}$. The landmarks comprise intensity quantiles, and minimum and maximum intensities. In a simplification of the original notation, we denote these all μ_i, with $i \in \{0, \ldots, N\}$, since the minimum and maximum intensities are also derived from quantiles. The target locations of these landmarks in the output histogram are denoted s_n; these can be derived from a training set, or an arbitrary target histogram.

For an input image $\mathcal{X}$ and output image $\mathcal{Y}$, the proposed transformation $\tau : x \mapsto y$ is defined

$$\begin{aligned} y &= \tau(x) \\ &= s_i + (x - \mu_i) \left(\frac{s_{i+1} - s_i}{\mu_{i+1} - \mu_i} \right), \qquad x \in [\mu_i, \mu_{i+1}] \end{aligned} \tag{8}$$

This transformation is illustrated in Figure 1. For simplicity, we assume the landmarks s are distributed equally in the target histogram – i.e. the target histogram is a uniform distribution, and

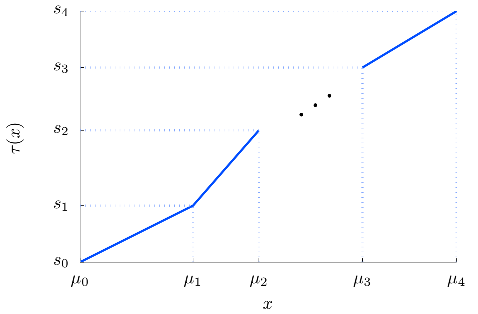


Fig. 1: Illustration of the Nyul et al. normalization, adapted from [2].

this approximates histogram *equalization*. As noted above, this assumption is not important for inter-image standardization. The output landmarks are therefore $s_i = \frac{i}{N}$, where $N+1$ is the number of quantiles. The transformation (8), now denoted τ_e, becomes

$$\tau_e(x) = \frac{i}{N} + (x - \mu_i)\left(\frac{1/N}{\mu_{i+1} - \mu_i}\right), \qquad x \in [\mu_i, \mu_{i+1}]. \tag{9}$$

Now, consider the functional representation of any quantile, $Q : \eta \mapsto x$, where η is a probability $\in [0,1]$. This can be defined mathematically as the functional inverse of the CDF of $\mathcal{X}$ [3],

$$Q(\eta) = P^{-1}(\eta). \tag{10}$$

If we define the probability $\eta = i/N$, and $\delta = 1/N$, then

$$\mu_i = Q\left(\frac{i}{N}\right) = Q(\eta) = P^{-1}(\eta) \tag{11}$$

$$\mu_{i+1} = P^{-1}(\eta + \delta). \tag{12}$$

Substituting this result into the denominator of the above expression, we obtain

$$\tau_e(x) = \eta + (x - \mu_i)\left(\frac{\delta}{Q(\eta+\delta) - Q(\eta)}\right), \qquad x \in [\mu_i, \mu_{i+1}]. \tag{13}$$

We notice the inverse of the Newton difference quotient for Q in the final term. If we assume small δ (large N), then this approximates $dQ(\eta)/d\eta = Q'(\eta)$,

$$\tau_e(x) \approx \eta + (x - \mu_i)\left(\frac{1}{Q'(\eta)}\right), \qquad x \in [\mu_i, \mu_{i+1}] \tag{14}$$

Next, we invoke the inverse function theorem [4], which states,

$$(f^{-1})'(f(x)) = \frac{1}{f'(x)}. \tag{15}$$

In our case, $f = Q$, $f^{-1} = Q^{-1} = P$, and $(f^{-1})' = P' = p$, which yields

$$\tau_e(x) \approx \eta + (x - \mu_i)\, p_x(Q(\eta)), \qquad x \in [\mu_i, \mu_{i+1}]. \tag{16}$$

Now, we exploit the recursive parameterization of τ_e, namely that $\eta = \tau_e(\mu_{i-1})$, with $\tau_e(\mu_0) = x_{\min}$. Substituting this result and replacing $\mu_i = Q(\eta)$, we obtain

$$\tau_e(x) \approx \tau_e(\mu_{i-1}) + (x - \mu_i)\, p_x(\mu_i), \qquad x \in [\mu_i, \mu_{i+1}]. \tag{17}$$

Rewriting the recursion as a cumulative sum, we have

$$\tau_e(x) \approx \sum_{j=0}^{i}\left[\left(\mu_{j+1} - \mu_j\right) p_x(\mu_j)\right] + \left(x - \mu_j\right) p_x(\mu_j), \quad x \in [\mu_i, \mu_{i+1}] \tag{18}$$

which we recognize as a non-uniform Riemann approximation of the cumulative integral of $p_x(x)$, with grid points $\{\mu_0, \ldots, \mu_N\}$. For large N, the last term is small, so it can be ignored,

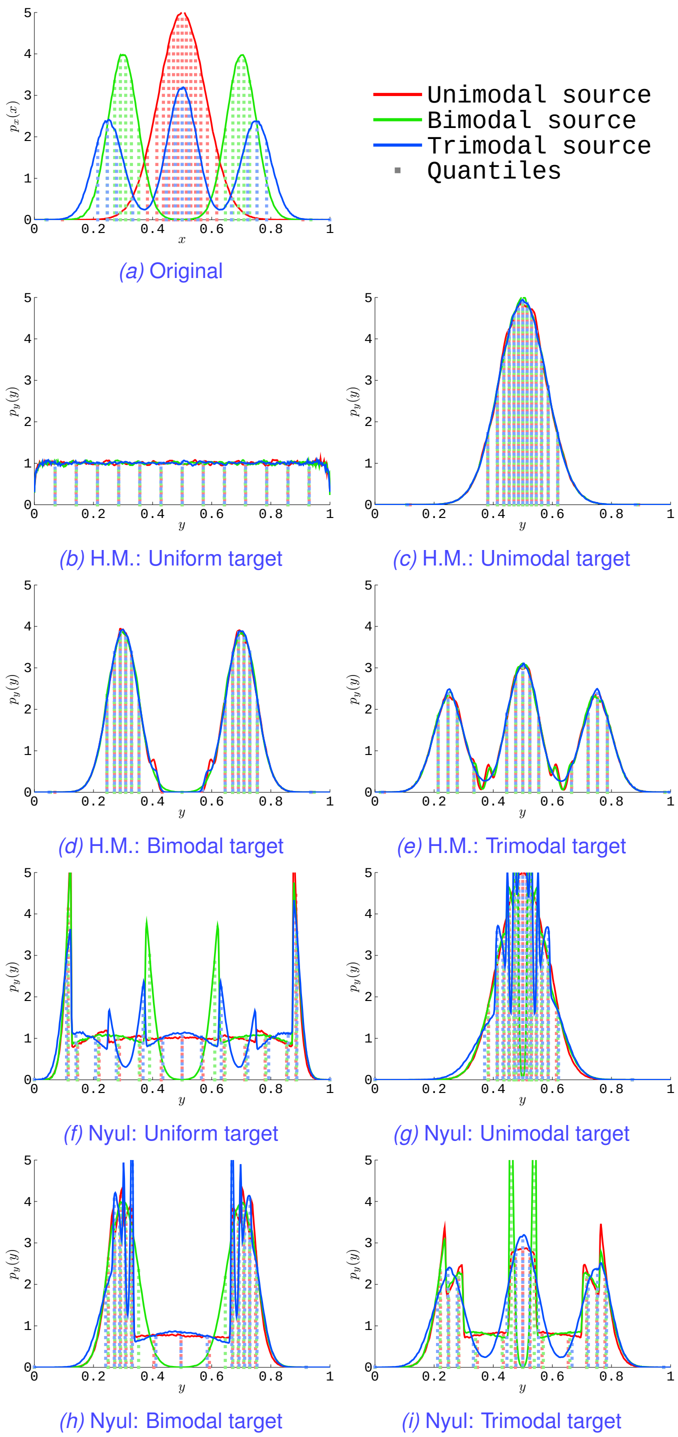
$$\tau_e(x) \approx \sum_{j=0}^{i}\left[\left(\mu_{j+1} - \mu_j\right) p_x(\mu_j)\right], \qquad x \in [\mu_i, \mu_{i+1}] \tag{19}$$

$$\approx P(x). \tag{20}$$

Therefore, for uniformly distributed output landmarks, the standardization method proposed by Nyul et al. in [1, 2] is a non-uniform trapezoidal Riemann approximation of the histogram equalization transform. For non-uniform output landmarks deriving from a target histogram, the composition property illustrated in (7) suggests how the Nyul method then approximates a histogram matching operation.

5 Experiments & Results

To validate these results, we first show that the choice of target histogram is unrelated to graylevel agreement between different images following ideal histogram matching. Three $100 \times 100 \times 100$ images are generated randomly to have unimodal, bimodal, and trimodal PMF, as shown in Figure 2a. Histogram matching is then used to match these images to one of four target histograms, including the uniform distribution – i.e. this is histogram equalization. The matched PMF are shown in Figure 2b–2e, including the image-specific quantiles. Following standardization, these quantiles agree almost perfectly, regardless of the target histogram: average absolute quantile difference is reduced to 2.76% of the original spread.



(a) Original

(b) H.M.: Uniform target

(c) H.M.: Unimodal target

(d) H.M.: Bimodal target

(e) H.M.: Trimodal target

(f) Nyul: Uniform target

(g) Nyul: Unimodal target

(h) Nyul: Bimodal target

(i) Nyul: Trimodal target

Fig. 2: Histogram matching of three source images to four different target histograms, using ideal histogram matching (H.M.) and the Nyul method ($N = 8$). Intensity quantiles are shown as vertical lines, with colour matching the image.

Next, we show that, under these same conditions, the Nyul method results in worse intensity agreement among distributions. Repeat-

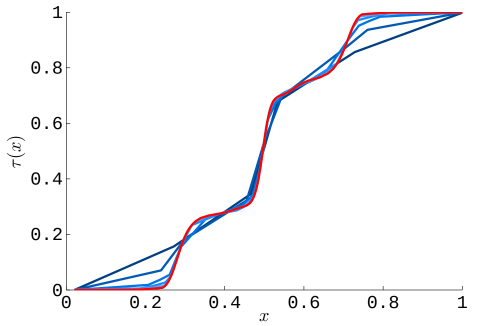
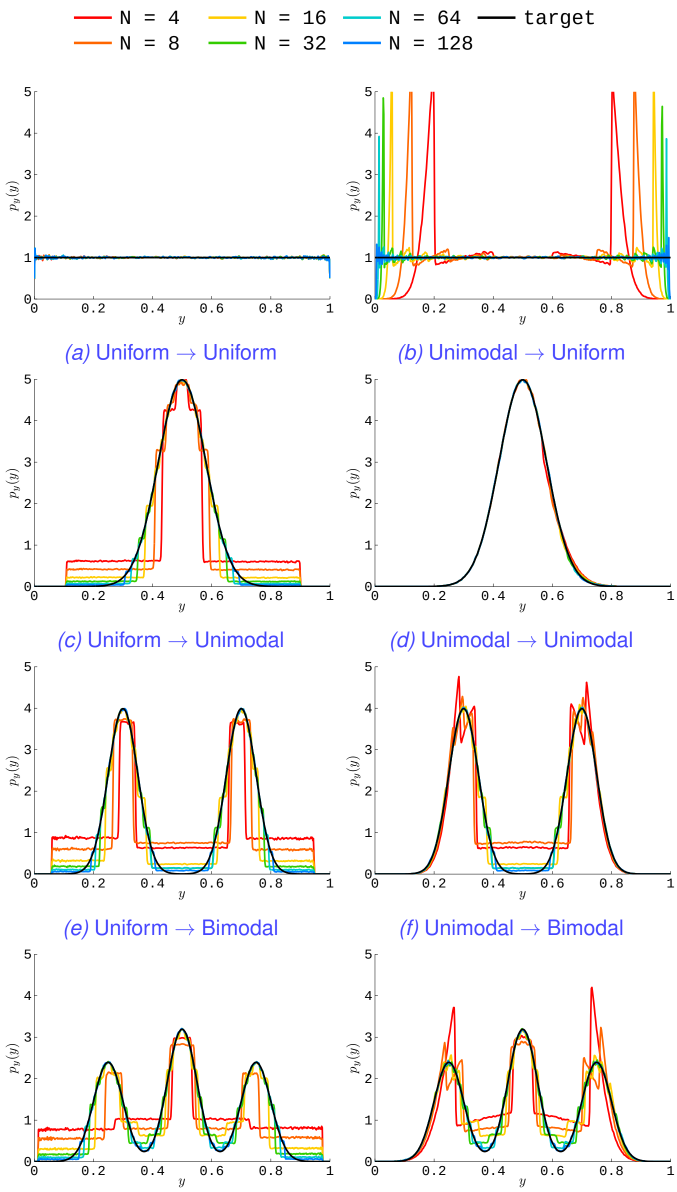


Fig. 3: Histogram equalization transformations: ideal (red) versus quantile-Riemann approximated (blue) with different N.

ing the above procedure using the Nyul with $N = 8$ method gives the results shown in Figures 2f–2i. This time, the average absolute quantile difference is reduced only to 6.76%.

As illustrated in Section 4, these differences could be expected to decrease if the Nyul method uses large N. To this end, we show that the Nyul transformation with uniform output landmarks converges to histogram equalization with increasing N. Using another random image of the same size, with trimodal PMF, the ideal equalizing transform $P(x)$ is computed as the CDF of the image intensities. The transformation $\tau(x) \approx P(x)$ from [1] is also computed using evenly spaced quantiles, with different $N \in \{2^2, \ldots, 2^7\}$. As shown in Figure 3, the approximation converges as N increases.

Finally, we explore the nature of the output histogram artifacts when using the Nyul method with small N. Two random images are generated, with uniform and unimodal PMF respectively. Using the Nyul method with $N \in \{2^2, \ldots, 2^7\}$, the intensities are matched to the same four target histograms as in Figure 2. The resulting PMF are shown in Figure 4. It can be seen that, in piecewise sections, the shape of the source PMF is maintained after standardization using this method. These artifacts are particularly noticeable in ranges of x where the source and target PMFs are most different .



(a) Uniform → Uniform
(b) Unimodal → Uniform
(c) Uniform → Unimodal
(d) Unimodal → Unimodal
(e) Uniform → Bimodal
(f) Unimodal → Bimodal
(g) Uniform → Trimodal
(h) Unimodal → Trimodal

Fig. 4: PMF of synthetic data following histogram matching using the Nyul method, with different N.

6 Conclusion

Histogram matching operations are popular in MR image analysis for intensity standardization. We have shown how a histogram matching operation can be decomposed into an image-specific histogram equalization and a second inverse transform to obtain the desired histogram. We note how this implies that the target histogram is unrelated to the objective of standardization across inputs. We also characterize the type of histogram matching approximation proposed by Nyul et al., show the resulting artifacts, and how this approximation is improved with increasing N.

References

[1] L. G. Nyul and J. K. Udupa. "On standardizing the MR image intensity scale". In: *Magn Reson Med* 42.6 (Dec. 1999), pp. 1072–1081.

[2] L. G. Nyul, J. K. Udupa, and X. Zhang. "New variants of a method of MRI scale standardization". In: *IEEE Trans Med Imaging* 19.2 (Feb. 2000), pp. 143–150.

[3] E. Parzen. *Quantile Functions, Convergence in Quantile, and Extreme Value Distribution Theory.* Tech. rep. Texas A and M Univ College Station Inst of Statistics, 1980.

[4] A. Nijenhuis. "Strong derivatives and inverse mappings". In: *The American Mathematical Monthly* 81.9 (1974), pp. 969–980.