IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 23, 2018, accepted April 19, 2018, date of publication May 4, 2018, date of current version June 5, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2833286

Preprocessing of Heteroscedastic Medical Images

PHILIP JORIS[1], WIM DEVELTER[2], WIM VAN DE VOORDE[2], PAUL SUETENS[1], FREDERIK MAES[1], DIRK VANDERMEULEN[1], AND PETER CLAES[1]

[1]Medical Imaging Research Center, Department ESAT-PSI, KU Leuven, 3000 Leuven, Belgium
[2]Department of Forensic Medicine, University Hospitals UZ Leuven, 3000 Leuven, Belgium

Corresponding author: Philip Joris (philip.joris@esat.kuleuven.be).

This work was supported in part by KU Leuven's Concerted Research Action GOA/11/006 and in part by UZ Leuven's KOF Mandate.

ABSTRACT Tissue intensity distributions in medical images can have varying degrees of statistical dispersion, which is referred to as heteroscedasticity. This can influence image contrast and gradients, but can also negatively affect the performance of general-purpose distance metrics. Numerous methods to preprocess heteroscedastic images have already been proposed, though most are application-specific and rely on either manual input or certain heuristics. We therefore propose a more general and data-driven approach that relies on the notion of intensity variance around each specific intensity value, simply referred to as intensity-specific variances. First, we introduce a method for estimating these variances from an image (or a collection of images) directly, which is followed by an illustration of how they can be used to define intensity-specific distance measures. Next, we evaluate the proposed concepts through various applications using both homo- and heteroscedastic CT and MR images. Finally, we present results from both qualitative and quantitative analyses that confirm the working of the proposed approaches, and support the presented concepts as valid and effective tools for (pre)processing heteroscedastic medical images.

INDEX TERMS Heteroscedastic, heteroscedasticity, image contrast, image enhancement, intensity-specific distributions.

I. INTRODUCTION

Tissue intensity distributions in medical images can have varying degrees of statistical dispersion. This phenomenon is known as heteroscedasticity, and influences image properties and processing methods for CT as well as MR images. In CT images, for example, narrowly distributed soft tissue intensities cover only a small part of the intensity axis. This reduces the dynamic range of these soft tissues and causes them to appear as low contrast objects. Furthermore, this leads to imbalanced image gradients, as edge magnitudes around broadly distributed bone intensities dominate edge magnitudes surrounding soft tissues. Finally, heteroscedasticity also negatively affects the performance of general-purpose metrics. The Euclidean distance, for example, fails to accurately depict intensity similarity when tissues are not identically distributed. To illustrate, a CT image intensity of 200 Hounsfield units (HU) is closer (and thus more similar) to 15HU than to 500HU. Yet we can distinguish a closer similarity between 200HU and 500HU as both correspond to bone tissue, whereas 15HHU corresponds to cerebrospinal fluid.

Despite these disadvantages, heteroscedasticity is not always taken into account or explicitly dealt with. In [1], for example, a framework is proposed for the segmentation of multiple organs in abdominal CT images. No preprocessing is used, even though low contrast soft tissue organs are targeted. Another example can be found in [2], where images are segmented into supervoxels before feeding them to a deep convolutional network. To obtain these supervoxels, however, the authors use SLIC [3], a supervoxel segmentation approach which is heavily affected by heteroscedasticity as it is based on k-means.

When images are preprocessed, it is most frequently in the form of a (non-)linear intensity transformation, masking and thresholding, or a combination of both. Intensity transformations are abundant (recent examples include [4]–[6]), though most methods are variants of Histogram Equalisation (HE) and Adaptive HE (AHE). Intensities are transformed on a global or local level with the aim of increasing contrast of certain structures such as brain tissues in head CT images [7] or fishbones on radiographs [8]. Masking and thresholding, on the other hand, are similar to the windowing of an image, whereby the goal is to include or exclude certain anatomical regions or tissues. How the required thresholds or masks are defined often varies between applications, with recent examples including manually seeded points in

combination with k-means to roughly estimate the liver position [9], and horizontal image scanning [10] or c-means [11] combined with morphological operations to obtain intracranial masks. Combinations of masking, thresholding, and intensity transformations are also common, with examples including [12]–[14].

Notwithstanding their widespread use, current preprocessing methods are application-dependent, heuristic, and do not incorporate tissue distribution widths explicitly. Thresholding an image requires threshold values which need to be defined manually or learned ad-hoc from the image or its histogram. The same goes for image masking, which calls for manually drawn masks or (a combination of) thresholds, clustering, and morphological operations. Contrast stretching, similarly, requires at least two thresholds to be defined. Finally, HE, AHE, and their variants rely on intensity frequencies to define an intensity transformation. Again, this is rather heuristic, and causes such methods to be biased towards the most frequent tissue classes, not per se towards the ones with the narrowest tissue distributions (and thus the lowest contrast).

We therefore propose a data-driven approach that requires no manual or ad-hoc tuning. The method relies on intensity-specific variances, which are estimated from an image (or a collection of images) directly to give an idea of the intensity variance for each intensity value. In section II-B, we introduce a method for estimating these variances, and illustrate how they can be used to define intensity-specific distance measures. Section II-C will then present several use cases of these variances and distances measures. More specifically, we discuss a global non-linear intensity transformation, image quantisation, and brain voxel classification, all of which are applied to homo- and heteroscedastic CT and MR images (accompanying MatLab code can be found at [15]). An overview of all used data is given in section II-A. Next, section III will present qualitative and quantitative results for all three applications, with the proposed concepts being compared to widely-accepted or state-of-the-art methods. Finally, section IV presents our final conclusions.

II. MATERIALS AND METHODS

A. DATA

Throughout this work, a variety of datasets was used to analyse and evaluate the methods discussed in section II-C. The data contained both CT and MR images, and is summarised in table 1. The first dataset comprised 72 head CT images [16], and is henceforth referred to as the head-CT dataset. All scans had an identical volume size and isotropic voxels. No ground truth segmentations were available. The second set consisted of 50 publicly available abdominal CT images from the "Multi-Atlas labeling beyond the cranial vault" MICCAI 2015 Challenge [17], simply referred to as the abdominal-CT data. Both volume and voxel sizes were variable. Ground truth segmentations were available for the spleen, kidneys, gallbladder, oesophagus, liver, stomach, aorta, inferior vena cava, portal and splenic vein, pancreas, and the adrenal glands. Two sets of 7 simulated MR images (T1 and T2 weighted) from the BrainWeb repository [18], [19] made up the third dataset. Each of the images was simulated using default ICBM parameters, with varying noise levels from 0% to 6% with a 1% increment. No bias field was added to the images, for reasons explained in section II-B. All images had identical volume sizes with isotropic voxels. Ground truth segmentations were available for background, CSF, gray matter (GM), white matter (WM), fat, muscle, skin, skull, and glial matter. Lastly, the Internet Brain Segmentation Repository dataset (ISBR V2.0) [20] was included, comprising 18 T1 weighted MR images. All scans had an identical volume size with variable anisotropic voxel sizes. Ground truth segmentations were again available for CSF, GM, and WM.

TABLE 1. Overview of the different datasets that are used throughout this work.

head-CT	Modality:	CT
	Region:	Head/neck
	Source:	Real
	Volume size:	271×271×239
	Voxel size:	1×1×1
	Ground truth:	No
Abdominal-CT	Modality:	CT
	Region:	Abdomen
	Source:	Real
	Volume size:	512×512×(85-198)
	Voxel size:	(0.54-0.98)×(0.54-0.98)×(2.5-5)
	Ground truth:	Yes (segmentations)
BrainWeb	Modality:	MRI (T1 & T2)
	Region:	Head
	Source:	Simulated
	Volume size:	181×217×81
	Voxel size:	1×1×1
	Ground truth:	Yes (segmentations)
IBSR	Modality:	MRI (T1)
	Region:	Head
	Source:	Real
	Volume size:	251 ×251×128
	Voxel size:	(0.9375-1)×(0.9375)×1.5
	Ground truth:	Yes (segmentations)

B. INTENSITY SPECIFIC DISTRIBUTIONS AND VARIANCES

In an attempt to develop a more data-driven approach for preprocessing heteroscedastic medical images, we made two simple assumptions about these images:

- Voxels of the same tissue type have similar intensities.
- Every voxel has at least one neighbouring voxel of the same tissue type.

The first assumption, also known as intensity homogeneity, applies to CT images as intensities are expressed using calibrated Hounsfield units. This assumption remains valid for individual MR images, provided that any bias field is removed. The second assumption states that any anatomical structure present in an image has a minimal volume of two voxels. This assumption holds true in most cases, if not all, considering today's imaging resolutions. Using these

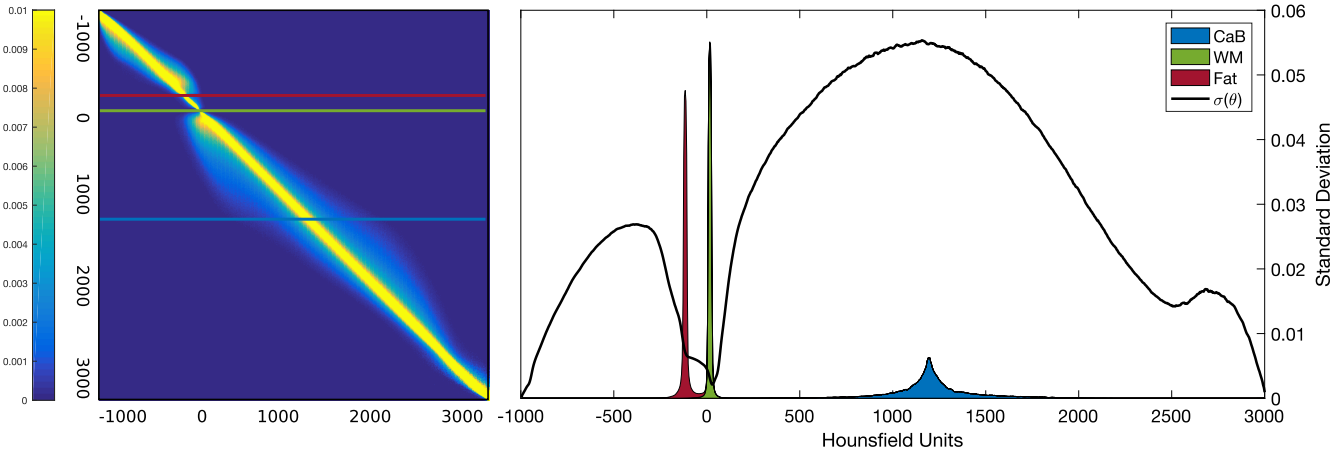

FIGURE 1. Example of intensity-specific distributions learned from a set of 30 head CT images using a total of 1000 bins. Each row in the left image is a probability distribution that integrates to one, representing the ISDs. Note how the distributions vary in between intensities. This is also illustrated on the right, were ISDs for -105HU, 25HU, and 1250HU are shown (corresponding to fat, WM, and CaB). The solid line plots the standard deviation for each ISD.

assumptions, we extend the idea of tissue heteroscedasticity to the level of individual intensities, and aim to assign a measure of statistical dispersion to each of an image's intensity values. We do this by first constructing intensity-specific distributions (or ISDs), which are obtained as follows: we start with an image $M = \{m_1, m_2, \ldots, m_n\}$ consisting of n voxels, assuming one out of b unique intensity values $V = \{v_1, v_2, \ldots, v_b\}$. Note that when intensities are expressed on a continuous instead of a discrete scale, intensities are first discretised using b bins. Additionally, discretisation can also be performed to reduce the number of intensities, and increase the number of samples per intensity value. This makes the estimation of ISDs more robust, and ensures that we can collect meaningful statistics of each ISD later on. To construct the ISD θ_i for intensity value v_i, we take all occurrences in M with an intensity value equal to v_i, and find the most similar neighbour of each of these occurrences using a 6-connectedness search. Similarity between neighbouring voxels is expressed as the absolute intensity difference. Next, we aggregate the intensities of all most similar neighbours, resulting in a intensity distribution that forms the ISD θ_i for v_i. Doing so for all elements in V results in $\Theta = \{\theta_1, \theta_2, \ldots, \theta_b\}$, constituting the full set of ISDs. Each of these ISDs can be seen as a non-parametric probability distribution, reflecting the local variation around each intensity value. Finally, we parametrise the ISDs' statistical dispersion using (1) and (2), where σ_i^2 and σ_i give the variance and standard deviation of the intensity value v_i.

$$\sigma_i^2 = \frac{\sum_{m \in \theta_i} (m - v_i)^2}{|\theta_i|} \tag{1}$$

$$\sigma_i = \sqrt{\sigma_i^2} \tag{2}$$

Note that the above assumes ISDs to be normally distributed. Other measures of spread such as the median absolute deviation or interquartile range are also possible, though we will use (1) and (2) throughout this work. In analogy with ISDs, we will refer to these variances as intensity-specific variances (or ISVs).

ISVs can be used directly in processing images, as the application in section II-C1 will illustrate. Alternatively, we can use ISVs to define similarity measures that incorporate these variances, allowing them to adapt to an image's content. Equation (3), for example, defines the similarity between two intensities v_i and v_j as the Euclidean distance between them, normalised by the product of their corresponding standard deviations, σ_i and σ_j. This distance measure satisfies symmetry, non-negativity and identity of indiscernibles. It does not, however, satisfy the triangle inequality, making it a semimetric. The applications discussed in sections II-C2 and II-C3 use this measure as a replacement for the standard Euclidean distance.

$$d(v_i, v_j) = \frac{(v_i - v_j)^2}{\sigma_i \sigma_j} \tag{3}$$

The reasoning behind the procedure outlined above is based on the assumptions made before. When looking at a single voxel, we assume it to have at least one neighbour of the same tissue type (assumption 2). To find this neighbour, we look for the surrounding voxel with the most similar intensity (assumption 1). The intensity difference between the original voxel and its most similar neighbour gives an idea of the variation in intensity for that specific intensity value. To get a more robust estimate of this variability, we aggregate the most similar neighbours from all voxels of that same intensity value (assumption 1), resulting in a single ISD. Note that when working with a calibrated intensity scale (such as HU in CT images), ISDs can be learned from a collection of images, as their intensities correspond. Fig. 1 shows the result of applying this procedure to a set of 30 images from the head-CT dataset. The continuous Hounsfield scale was discretised using a total of 1000 bins, which resulted in 1000 ISDs. The left side of Fig. 1 shows these ISDs, where each row represents a single distribution, normalised

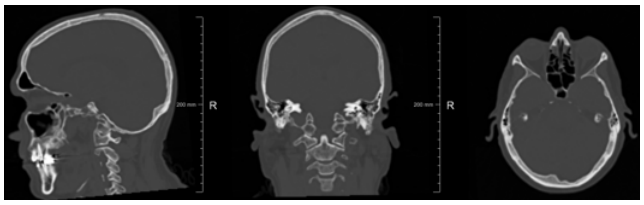

FIGURE 2. Example of a standard head CT image. Note how bone tissues take up most of the image's dynamic range, while soft tissues are barely discernible.

to integrate to one. On the right side, individual ISDs are plotted for -105HU, 25HU and 1250HU, which are intensities corresponding to fat, WM, and cancellous bone (CaB). The solid black line plots the standard deviation for each ISD shown on the left. As can be seen, fat has a higher local variation than white matter, whereas CaB shows the largest level of dispersion. These differences in distribution widths are in accordance with the problems discussed in section I (e.g. bone tissue having the widest distribution), and supports the above method for estimating these local variances.

C. APPLICATIONS

1) CONTRAST REDISTRIBUTION

As discussed before, CaB and cortical bone (CoB) comprise a large part of a CT image's intensity axis, thus taking up most of the image's dynamic range. This results in poor overall contrast, which is apparent when looking at Fig. 2. CaB and CoB appear as high contrast objects in the image, whereas individual soft tissues such as CSF, fat, and muscle remain mostly indiscernible. We therefore propose a method for contrast enhancement that redistributes intensities along the intensity axis by relying on an image's ISVs. Equation (4) defines a non-linear transformation, whereby the i^{th} original intensity value v_i is transformed to its new value v_i^t by taking the cumulative sum of the reciprocals of the ISVs up to and including σ_i. The remainder of (4) uses the sum of all ISVs and v_b and v_1 (the maximum and minimum intensity respectively) to ensure that the output image has the same intensity range as the input image. This transformation assumes ISDs to be normally distributed and consequently characterises their statistical dispersion using the standard deviation. By dividing the distribution of each bin by their standard deviation, we obtain normal distributions with unit standard deviation, effectively removing differences in statistical dispersion between distributions.

$$v_i^t = T(v_i) = \sum_{j=1}^{i} \frac{1}{\sigma_j} \cdot \frac{(v_b - v_1)}{\sum_{k=1}^{b} \frac{1}{\sigma_k}} + v_1 \tag{4}$$

Fig. 3 shows an example transformation that was obtained by using the standard deviations from Fig. 1. Note how brain tissue intensities (indicated in blue) are expanded by the transformation, as their ISVs are lower relative to other intensities. Conversely, bone tissue intensities (indicated in red) are compressed, as their ISVs are the largest.

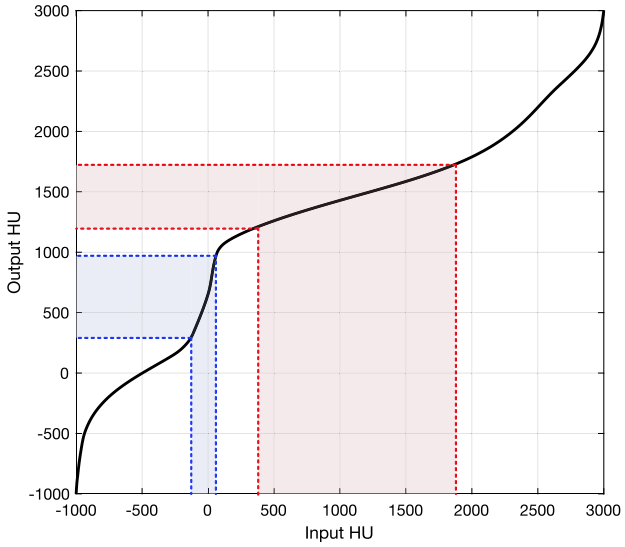

FIGURE 3. Example of a non-linear transformation to enhance image contrast. The transformation was obtained using (4) and the ISVs from Fig. 1.

We analysed the proposed transformation and evaluated its effect on image contrast, image gradients, and edge detectors using 20 images from the abdominal-CT dataset. Each image was discretised into 1000 distinct intensities, after which ISDs were calculated. Note that ISDs were learned from the full set of images in this case, since CT images use calibrated intensities. ISVs were then calculated and used to transform the 20 images with the transformation given in (4). Three classical gradient based edge detectors (the finite central difference, the Prewitt [21], and the Sobel and Feldman [22] operators) were then used to calculate gradient magnitude maps, which were thresholded to classify voxels into edges and non-edges. The threshold value was varied while resulting labels (edge or no edge) were compared to a ground truth, resulting in receiver operating characteristic (ROC) curves. Ground truth edges were derived from the available ground truth segmentations using an approach similar to [23]. Results are discussed in section III-A.

2) IMAGE QUANTISATION

Image quantisation aims to reduce the number of intensity values in an image, while at the same time maintaining close resemblance to the original. Two well-known approaches are interval discretisation (ID), or uniform quantisation, and frequency discretisation (FD), or quantile quantisation. The former divides the intensity axis into equally sized bins with varying sample counts per bin, whereas the latter divides the image intensities into varying sized bins with equal sample counts per bin. Both methods replace each input intensity by the centre value of the bin that intensity falls within. Ideally, we want to ensure that different tissues remain distinguishable after discretisation, and that each tissue is approximated with the same degree of accuracy. The aforementioned methods, however, fail to do so, as both rely on simple heuristics for partitioning the intensity axis.

We therefore propose a new quantisation method based on the concept of ISDs, enabling the discretisation to handle tissue or intensity heteroscedasticity. Quantisation is achieved by using the unweighted pair group method with arithmetic mean (UPGMA) [24], which is an agglomerative clustering approach whereby all intensities are initially treated as separate clusters and iteratively combined until a specified number of clusters is reached. At each iteration, the two closest clusters are merged to create a new single cluster. The distance between any two clusters C_A and C_B is taken to be the average of all pairwise distances between all elements of these clusters, as given by (5).

$$D_{UPGMA}(C_A, C_B) = \frac{1}{|C_A||C_B|} \sum_{i \in C_A} \sum_{j \in C_B} d(i,j) \tag{5}$$

In (5), the distance $d(i,j)$ between intensities i and j is given by the similarity measure from (3). When the required number of clusters is reached, each intensity value in the original image is replaced by the centre value of the bin that intensity falls in, as with uniform and quantile discretisation. The proposed hierarchical clustering quantisation method will be referred to as HCQ.

All three methods (ID, FD, and HCQ) were analysed and compared in terms of their quantisation error using 20 images from the head-CT dataset and all BrainWeb T1 and T2 images. Each of the images was first discretised using 1000 bins for CT and 200 for MR images, after which ISDs were learned for each image separately. Note that ISDs could have been learned from the full set of CT images, though we refrained from this to keep a fair comparison (ID and FD only use the information from a single image). Each image was then quantised into a number of bins (ranging from 2 to 120), after which the mean squared error (MSE)(6) was used to evaluate the quantisation error between the original (M) and quantised image (M^q). Equation (6) expresses overall approximation quality, but gives no indication of the relative approximation error between different tissue classes. We therefore divided each quantised image into k disjoint sets of voxels $M = \{S_1, S_2, \ldots, S_k\}$ using the available ground truth segmentations, and calculated the normalised mean squared error (NMSE) for each tissue separately using (7). NMSE values are scaled by the standard deviation of each tissue class, meaning a comparison between tissues and images can be drawn. Results are discussed in section III-B.

$$MSE(M, M^q) = \frac{1}{n} \sum_{j=1}^{n} (M_j - M_j^q)^2 \tag{6}$$

$$NMSE(S_i, S_i^q) = \frac{MSE(S_i, S_i^q)}{\sigma(S_i)} \tag{7}$$

3) VOXEL CLASSIFICATION

Voxel classification attempts to divide an image's voxels into coherent groups that follow the natural structure of the image. A popular example is k-means clustering [25], a method that minimises the within-cluster sum of squares using the Euclidean distance metric. Because of its objective function (8), k-means will generate clusters of similar spatial extent in the feature space, implicitly assuming the input data to be roughly homoscedastic. To adapt k-means towards heteroscedastic data, we changed the cost function of (8) to that of (9), where the Euclidean distance was replaced by the similarity measure proposed in (3). To optimise this new cost function, we employ an efficient histogram-based optimisation procedure. We start by building a standard histogram using b bins, and randomly assign one out of k class labels to each bin, resulting in k sets of intensities $S = \{S_1, S_2, \ldots, S_k\}$. The initial labels are then iteratively updated in order to minimise the cost function in (9). The complete procedure is described in algorithm 1, and is henceforth referred to as ISDClstr.

$$\underset{S}{\operatorname{argmin}} \sum_{j}^{k} \sum_{a \in S_j} \sum_{b \in S_j} (a-b)^2 \tag{8}$$

$$\underset{S}{\operatorname{argmin}} \sum_{j}^{k} \sum_{a \in S_j} \sum_{b \in S_j} d(a,b) \tag{9}$$

We analysed the performance of the proposed clustering procedure and standard k-means using both CT and MR images. First, we compared both approaches qualitatively by classifying images from the head-CT dataset. Second, we evaluated both methods quantitatively by applying them to the problem of brain voxel classification in MR images. We also included results from SPM12 [26] (Statistical Parametric Mapping), a publicly available software suite that is frequently used in neurological studies. Because of SPM's widespread use and maturity, we use it as a reference. Each image (BrainWeb T1, BrainWeb T2, and IBSR T1) was first processed to extract the brain, after which the remaining voxels were classified into CSF, GM, and WM using k-means, ISDClstr, and SPM. Note that ISVs were learned separately for each image, as MR images generally lack intensity standardisation. Finally, segmentation accuracy was quantified using the Dice similarity coefficient. Results are discussed in section III-C.

III. RESULTS AND DISCUSSION

A. CONTRAST REDISTRIBUTION

Fig. 4 shows the head CT from Fig. 2 after applying the transformation from Fig. 3. The transformation used the image's ISVs to redistribute intensities and increase overall image contrast. Soft tissues take up a larger part of the dynamic range than before, enhancing their visibility. Note how CSF, fat, and muscle mass are distinguishable in Fig. 4, whereas this is not the case in Fig. 2. Bone intensities, on the other hand, have a reduced dynamic range, but retain sufficient contrast. A similar example is presented in Fig. 5, where an abdominal CT is shown before and after applying the proposed transformation. Comparing Fig. 5(a) and Fig. 5(b), there is again a clear difference in contrast, whereby soft tissues have an increased dynamic range at the cost of that

Algorithm 1 ISDClstr

Precondition: X is a vector contain b intensity values; H is a vector also of length b, containing sample counts per intensity in X; k specifies the number of clusters.

function Cluster(X, H, k)
  $L \leftarrow Initialise$ ▷ Initialise labels
  $c \leftarrow Cost(X, H, L)$
  while *not_converged* **do**
    for $i \leftarrow 1$ to b **do**
      for $j \leftarrow 1$ to k **do**
        $l_{old} \leftarrow L(i)$
        $L(i) \leftarrow j$
        $c_{new} \leftarrow Cost(X, H, L)$
        if $c_{new} < c$ **then**
          $c \leftarrow c_{new}$
        else
          $L(i) \leftarrow l_{old}$ ▷ Reset label
        end if
      end for
    end for
  end while
  return L
end function
function Cost(X, H, L) ▷ Calculate cost
  $c \leftarrow 0$
  for $i \leftarrow 1$ to b **do**
    for $j \leftarrow 1$ to b **do**
      if $L(i) = L(j)$ **then**
        $c \leftarrow c + H(i)H(j)d(X(i), X(j))$ ▷ (3)
      end if
    end for
  end for
  return c
end function

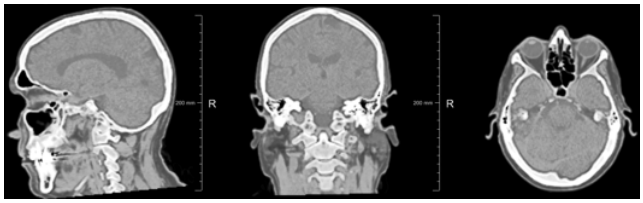

FIGURE 4. Same head CT as in Fig. 2, after applying the transformation shown in Fig. 3. Note how overall contrast is increased, and various soft tissues such as CSF, fat, and muscle mass become distinguishable.

of bone tissue. This trade-off is also visible in the images' histograms, shown in Fig. 5(c) and Fig. 5(d). Prior to the transformation, most soft tissues are contained in the narrow peak around 0HU, whereas they range from 700HU to 2000HU after. The opposite is true for bone tissue intensities that initially range from 700HU to 3000HU, but span a much smaller range after the transformation. This redistribution of intensities clearly changes an image's contrast, but also

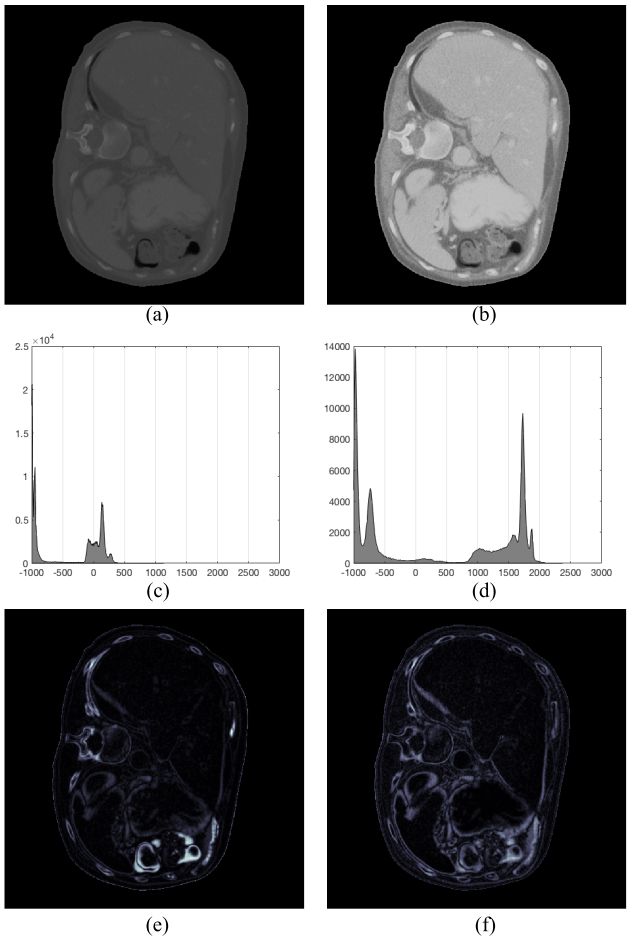

FIGURE 5. Axial slice of an abdominal CT before (a) and after (b) applying the transformation proposed in section II-C1. Corresponding histograms are shown in (c) and (d), with resulting Sobel gradient magnitude maps in (e) and (f).

affects its gradients, as can be seen in the gradient magnitude maps in Fig. 5(e) and Fig. 5(f). Before the transformation, magnitudes are dominated by interfaces between soft tissues and bone or air. Edges between soft tissues themselves have a much lower gradient magnitude, as a result of their narrow intensity distributions. When looking at the magnitude map in Fig. 5(f), interfaces between various tissues have more equal magnitudes. Edges around the ribs, for example, have a lower magnitude than before, while edges between the spleen, stomach, and liver have an increased gradient magnitude. Logically, balancing the image contrast has also led to balanced gradient magnitudes, which can be a desirable characteristic. Taking the example of edge detection, imbalanced magnitudes will make it difficult to select a single threshold to classify voxels as edges and non-edges. When magnitudes are less divergent, choosing a single threshold becomes more straightforward. This is confirmed by the ROC-curves shown in Fig. 6, which result from classifying the gradient magnitude maps from 20 images using a variable threshold. Note that each ROC-curve shifts towards the top-left corner of the plot after transforming the images, indicating that the

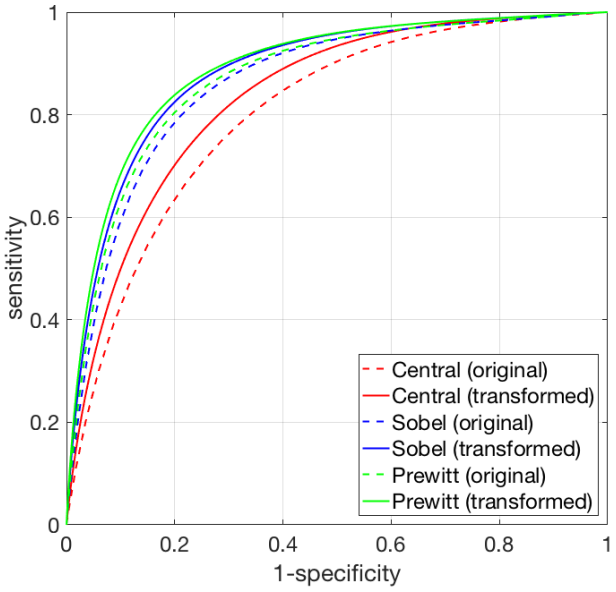

FIGURE 6. ROC-curves obtained by thresholding gradient magnitude maps of three different edge detectors. Note how performance increased by applying the proposed transformation.

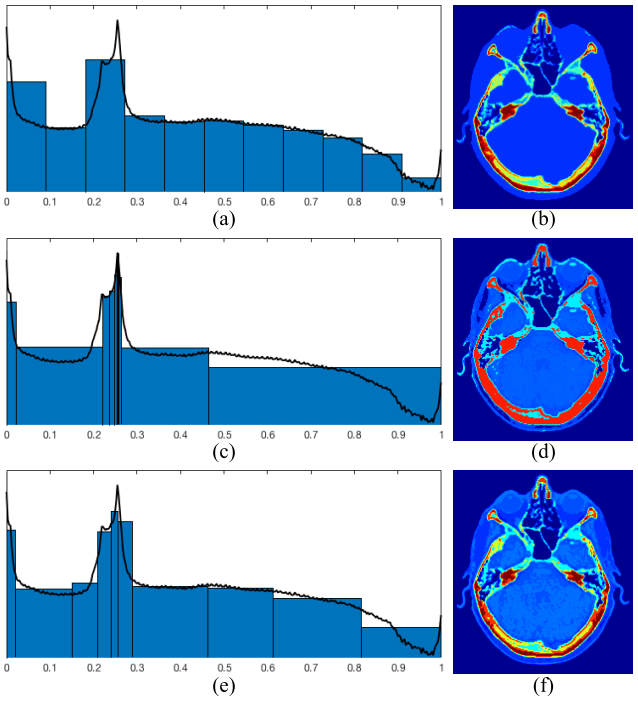

FIGURE 7. Log histograms of a head CT image after quantisation using ID (a), FD (c), and HCQ (e) using 11 bins. Figures (b), (d), and (f) show an arbitrary slice of the quantised image.

classification performance increased by preprocessing the images with the proposed transformation.

B. IMAGE QUANTISATION

Section II-C2 discussed a new quantisation method, alongside interval and frequency discretisation. Fig. 7 shows a qualitative comparison of these methods, showing the log histogram and an arbitrary slice from a quantised head CT image for each method. Interval discretisation partitions the intensity axis into equally sized bins, spreading the quantisation error roughly equally across all intensities. This means that, under the assumption of homoscedasticity, the quantisation error is spread equally across tissues as well. This is, however, not the case when dealing with heteroscedastic images, as can be seen in Fig. 7(a). ID overfits on bone tissue, using many bins to quantise its range of intensities, whereas it underfits on soft tissues, using only a single bin to approximate all soft tissues combined. This leads to a quantised image (Fig. 7(b)) wherein a relatively high level of detail is present in bone tissues, while almost no detail is left on soft tissues. Frequency discretisation, on the other hand, quantises an image using variable bin widths by relying on intensity frequencies (Fig. 7(c)). The majority of an image's voxels will therefore have a low quantisation error, while larger errors are tolerated on the less frequent intensities. Soft tissues are the most prevalent in a head CT, leading to small bins being used to approximate them, while bone tissues are approximated using larger bins as they appear most frequently. Logically, this causes FD to overfit on soft tissues, showing a relatively high level of detail, and underfit on bone tissues, displaying a relatively low level of detail (Fig. 7(d)). The proposed HCQ method, finally, partitions the intensity axis by relying on ISVs, permitting the use of variable bin widths without being biased by tissue frequencies. Smaller bins are used to approximate intensities with relatively small ISVs, while larger bins are used to approximate intensities with relatively large ISVs, as seen in Fig. 7(e). Several bins are used to approximate both soft and bone tissues, resulting in a quantised image (Fig. 7(f)) that has a balanced overall detail.

To evaluate the overall approximation accuracy, we repeated the above quantisation for 20 subjects of the head-CT dataset using various quantisation levels, and calculated the mean squared error between the original and the quantised images. Results are shown in Fig. 8, from which it is clear that FD has the largest quantisation error. FD underfits the broadly distributed bone tissue intensities, leading to large MSE values. ID, on the other hand, overfits on bone tissue, decreasing the quantisation error. However, relatively few bins were used to approximate the overwhelming amount of soft tissue voxels, consequently leading to a large amount of (albeit small) quantisation errors on soft tissues. HCQ, finally, balanced the partitioning of the intensity axis across the image's tissues, resulting in the lowest overall quantisation error. Note that this would not always have been the case. For example, when the frequency of bone tissue voxels would have been higher, ID would have had a lower overall MSE, as it strongly overfits the broadly distributed bone tissue intensities.

Fig. 7 and Fig. 8 suggest that HCQ, in contrast to ID and FD, is capable of approximating an image's tissues with the same relative accuracy. To assess this, we quantised all T1 and T2 weighted MR images from the BrainWeb dataset using

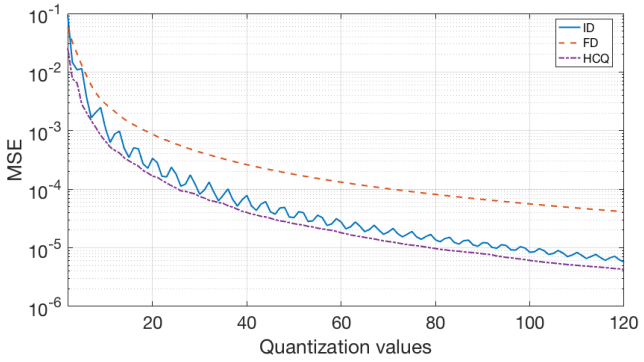

FIGURE 8. Average overall MSE for ID, FD, and HCQ. Quantisation levels vary from 2 to 120. Note the logarithmic scale.

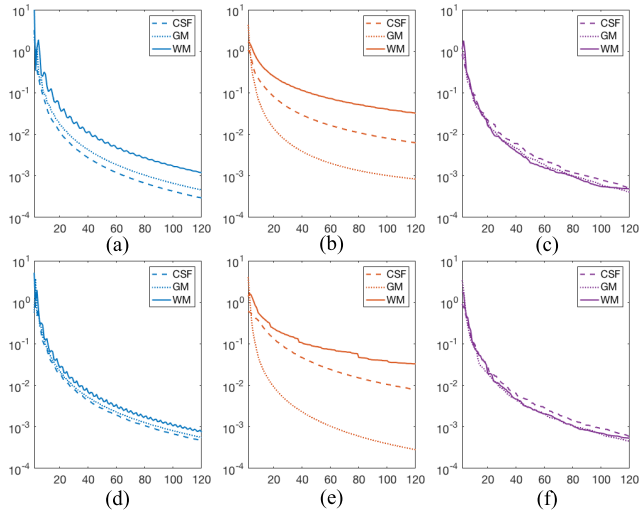

FIGURE 9. Average NMSE of 7 T2 weighted images for ID, FD, and HCQ (left to right).

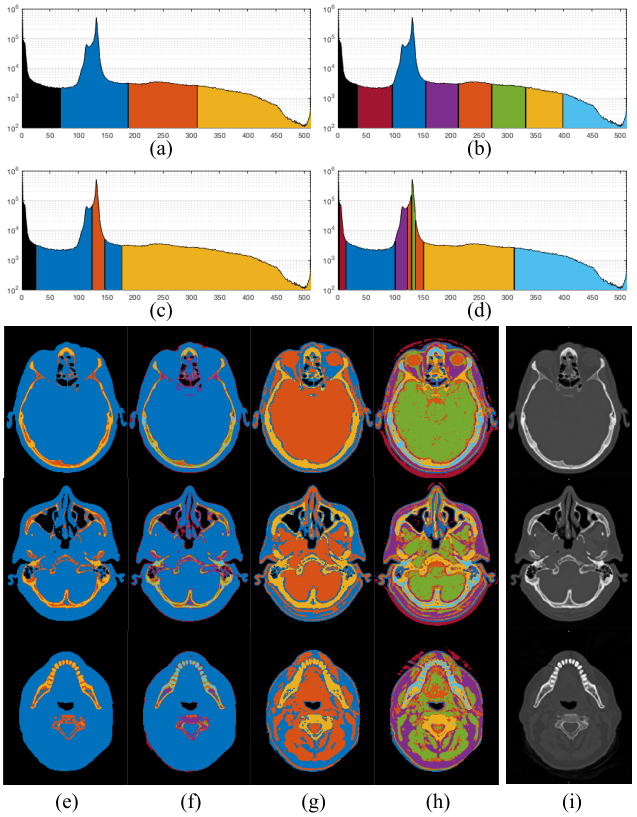

FIGURE 10. Log histograms from a head CT that was grouped into 4 and 8 clusters using k-means (a,b) and ISDClstr (c,d). Corresponding slices are shown in (e,f) for k-means and (g,h) for ISDClstr. Original slices are shown in (i).

various quantisation levels, and calculated the normalised MSE for CSF, GM and WM separately. When tissues are quantised with the same relative level of detail, curves will coincide. Conversely, when a method over- or underfits a certain tissue, differences will occur. Figs. 9(a)-(c) show results for the T2 images, which showed notable heteroscedasticity with average tissue standard deviations of 0.53, 0.34, and 0.14 for CSF, GM and WM respectively. This caused ID to overfit on CSF (having the broadest distribution) and underfit on WM (having the narrowest distribution). Logically, CSF and WM showed varying quantisation errors (Fig.9(a)). Even more divergent NMSE values were found for FD, strongly overfitting GM intensities since they were the most prevalent tissue type (46%). Consequently, GM voxels show a much lower quantisation error compared to WM and CSF (Fig. 9(b)). HCQ curves, on the other hand, are notably more similar compared to those of ID and FD, indicating that quantisation errors were indeed more equally spread across different tissues (Fig. 9(c)). To quantify the divergence of the curves, we calculated the standard deviation of the NMSE values for all tissues combined per level. On average, ID curves were 7.22 times more divergent than those of HCQ, while FD curves were 123 more divergent. Similar results for T1 images are shown in Figs. 9(d)-(f), though T1 tissue distributions showed less heteroscedasticity, with average standard deviations of 0.40, 0.34 and 0.25 for CSF, GM and WM respectively. On average, ID and FD curves respectively were 1.7 and 92 times more divergent compared to HCQ values.

C. VOXEL CLASSIFICATION

Section II-C3 discussed a new clustering method, which was based on the concept of ISVs. Fig. 10 shows a qualitative comparison of ISDClstr and standard k-means, where the intensities of a head CT were clustered into four and eight groups. Referring to the four-cluster result of k-means in Fig. 10(a), note how clusters have a similar spatial extent, which is a direct result of the k-means objective function. Corresponding image slices are shown in Fig. 10(e). Two clusters are used to represent bone intensities, whereas one cluster combines all soft tissues and another roughly corresponds to air. Although this grouping could be considered a logical partitioning of the image's intensities, this is likely more by accident than design. This becomes apparent when looking at Fig. 10(b) and Fig. 10(f). Five out of eight clusters are used to model bone intensities, whereas all soft tissues

TABLE 2. Dice coefficients for SPM, k-means and ISDClstr on T1 BrainWeb data. Highest scores are in bold. Percentages represent the noise level.

		0%	1%	2%	3%	4%	5%	6%
CSF	SPM	83.2	82.7	82.0	82.1	81.9	80.4	79.7
	k-means	**89.1**	**89.0**	**88.6**	**88.1**	**87.2**	**85.3**	**83.6**
	ISDClstr	88.0	87.9	86.8	86.9	85.2	82.6	81.0
GM	SPM	80.4	86.3	86.4	86.0	**85.2**	**83.8**	**82.6**
	k-means	**90.1**	**89.9**	88.4	**87.3**	84.5	81.4	77.5
	ISDClstr	88.3	89.0	**87.0**	86.4	83.1	78.9	74.9
WM	SPM	72.3	89.6	**91.4**	**91.2**	**90.1**	**88.5**	**87.4**
	k-means	**93.2**	**92.9**	91.2	90.4	87.6	84.7	81.2
	ISDClstr	90.8	92.8	91.1	90.5	87.8	84.7	81.2

TABLE 3. Dice coefficients for SPM, k-means and ISDClstr on T2 BrainWeb data. Highest scores are in bold. Percentages represent the noise level.

		0%	1%	2%	3%	4%	5%	6%
CSF	SPM	80.8	89.1	**90.8**	**90.0**	**88.6**	**87.2**	**85.8**
	k-means	64.6	64.7	64.9	66.0	77.0	76.2	75.1
	ISDClstr	**91.7**	**91.9**	90.7	88.0	84.1	80.7	77.6
GM	SPM	80.7	83.4	83.7	83.3	**82.7**	**81.7**	**80.8**
	k-means	27.2	27.4	28.4	32.9	66.4	65.9	65.2
	ISDClstr	**88.3**	**88.0**	**86.8**	**84.2**	80.3	76.9	73.7
WM	SPM	82.7	81.6	81.8	82.6	83.6	84.2	85.1
	k-means	81.2	81.2	81.3	82.0	85.0	84.8	84.3
	ISDClstr	**86.6**	**86.5**	**86.4**	**86.2**	**86.1**	**85.9**	**85.5**

remain together in a single cluster. The last two groups model air and mostly partial volume voxels. K-means is affected in a similar way as frequency discretisation before, meaning it tends to overfit broadly distributed bone tissue while underfitting the narrow distributions of soft tissues. ISDClstr, on the other hand, creates clusters that have highly variable spatial extents, as seen in Fig. 10(c) and Fig. 10(d). Clusters can clearly assume different sizes, and do not even need to form connected sets of intensity values. This is a result of minimising the semimetric from (3), and can be seen for example in the blue cluster of Fig. 10(c). Furthermore, clusters also appear to model certain tissues in the image as a result of this semimetric. In Fig. 10(g), for example, clusters correspond to air, a combination of fat and some partial volume intensities, a combination of various soft tissues such as WM, GM, CSF, and muscle mass, and finally a combination of CaB and CoB. The same holds for the eight-cluster result in Fig. 10(h). The purple cluster, for example, corresponds to fat tissue, while the orange cluster combines CSF with intraocular fluid. The green cluster combines brain and muscle mass, as their intensities are very similar. The yellow and light blue clusters, finally, correspond to CaB and CoB bone. These groupings are vastly different from those of k-means, and are simply the result of changing the metric that was minimised.

Above results, though qualitative, indicate that ISDClstr is capable of implicitly including image and tissue-specific information through the use of the proposed semimetric. To objectively quantify this, we turned to the problem of brain voxel classification of MR images using k-means and ISDClstr. Results for SPM were also included for reference. Tables 2 and 3 show results for the BrainWeb T1 and T2 images respectively. The T1 images are fairly homoscedastic, as previously mentioned, resulting in only small differences between k-means, ISDClstr and SPM. The T2 images, on the other hand, showed larger heteroscedasticity. As a result, k-means clearly failed to find correct classifications for GM and CSF, leading to the low Dice scores shown in table 3. ISDClstr, on the contrary, achieved Dice scores similar to those of SPM, and surprisingly even outperforms SPM for images with low levels of noise. This is likely caused by SPM's assumption that tissue distributions can be approximated by a Gaussian Mixture Model. Finally, Fig. 11 shows Dice scores for the IBSR dataset. Tissues

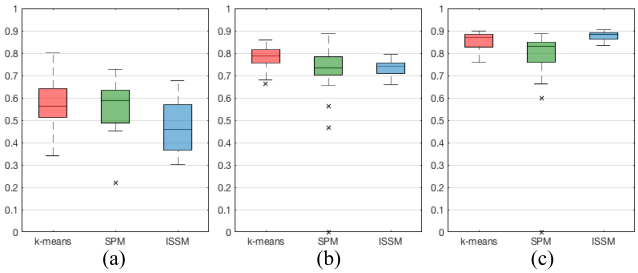

FIGURE 11. Dice coefficients for IBSR images using k-means (red, left), SPM (green, centre) and ISDClstr (blue, right).

showed little heteroscedasticity, with average standard deviations of 0.36, 0.38 and 0.26 for CSF, GM and WM respectively. Overall Dice scores are lower, though all three methods perform similarly.

D. DISCUSSION

Above results illustrate how heteroscedasticity can affect image properties and image processing methods. They also confirm that the proposed intensity-specific distributions and similarity measure can be used to better handle such heteroscedastic data. The implications of these results however, go beyond the discussed applications.

The intensity rescaling from section III-A redistributes intensities with the aim of increasing overall contrast, much like the popular histogram equalisation (HE) [27] and adaptive histogram equalisation (AHE) [28] methods. Such transformations deliver a visual advantage, but are also a popular preprocessing step as they can tune image quality towards a specific application [29] and improve the distinction of image features in image classification tasks [30]. The proposed transformation, however, differs from HE and AHE in the way it defines its transformation. HE and AHE use intensity frequencies, meaning that they increase contrast on the most frequent intensities at the cost of the less frequent ones. This is similar to the over- and underfitting of frequency discretisation (see section III-B). The proposed transformation is not biased by intensity imbalances as it uses ISVs, making it a suitable preprocessing step for various applications. Intensity clustering methods, for example, may benefit from ISV-based intensity normalisation, since tissue distributions become homoscedastic and thus easier to group. This was

already shown in section III-C, where k-means failed to group intensities correctly when the data was heteroscedastic. CT image registration may also benefit since similarity metrics such as the sum of squared or absolute differences are often used. Such metrics are biased towards aligning bone intensities at the cost of soft tissues, simply because bone intensities have a higher misalignment cost due to their wider distribution. Analogous problems may also arise in patch-based frameworks as the same metrics are often used to search for similar patches. Prior normalisation of the intensities could remove this bias. The proposed transformation also corrects imbalanced gradient magnitudes, which was discussed in section III-A. This in turn, can be advantageous for other applications, since gradients and edges are often used as an attracting force in cost functions in shape analysis. Examples can be found in image registration [31], [32], and image segmentation relying on for example watersheds [33], level-sets [34], [35], or model based approaches such as snakes [35] and Active Shape Models [36].

Image quantisation reduces the number of unique intensity values in an image and is therefore sometimes used in lossy image compression [37]. Using the proposed hierarchical approach, a minimum number of bins can be used for a given degree of approximation accuracy, while the quantisation error is spread equally across various tissues. This strongly contrasts the over- and under-fitting of ID and FD. Further, quantisation is closely related to constructing histograms, which are frequently used in image processing. Multimodel image registration using mutual information [38], [39], for example, uses image histograms. The proposed approach allows to build histograms that implicitly follow the structure of the image and its tissues, while at the same time having a small number of bins. This reduces computational complexity, but also ensures that different tissues remain discernible, even with a low number of bins. In addition, quantisation simply reduces the complexity of an image, as fewer intensities are used to represent it. This can increase efficiency in methods such as the histogram-based clustering outlined algorithm 1, or reduce requirements to store and process images, which is an active research topic in deep learning [40]. When intensities are used as random variables, quantisation also reduces the state space, meaning that further processing steps, such as intensity-based machine learning algorithms or Markov Random Fields, may benefit as well.

Finally, results in section III-C illustrated how the proposed semimetric can enable algorithms to implicitly include tissue-specific information and cope with heteroscedastic tissue distributions. The proposed clustering method displayed competitive performance compated SPM, though the latter obtained higher Dice scores for noisier images. This is expected, since SPM enables soft-clustering by using a Gaussian Mixture Model (GMM) that allows for overlapping clusters. Additionally, this GMM is combined with a Markov Random Field for spatial consistency, further increasing SPM's ability to label noisy voxels correctly. ISDClstr, on the other hand, is far more simple, and less time-consuming than SPM. It can therefore be used as an efficient and robust initialisation for more complex methods (such as GMMs), or in itself be used as a fast, simple, yet accurate approach, without the need for image registration or initial cluster estimates. The approach can also be extended towards multimodel intensity clustering, or be combined with more complex approaches such as soft-clustering [41] or Markov Random Fields [42] to enforce spatial constraints on the voxel labels. More importantly, however, the proposed similarity measure is easy to implement and can readily be integrated into existing applications and frameworks. Any algorithm that requires a metric for intensity similarities can use the proposed semimetric and enjoy its benefits.

IV. CONCLUSION

Throughout this work, we showed that medical images can have heteroscedastic tissue intensity distributions, and argued how this can have an affect on standard and widely-accepted image processing methods. We demonstrated this deficiency extensively through the applications of edge detection, image quantisation, and voxel classification, using standard methods to process both homo- and heteroscedastic CT and MR (T1 and T2) images. We also argued that existing preprocessing methods, if used, are often heuristic and tuned towards specific applications. In an attempt to handle heteroscedastic data more adequately, we proposed the concepts of intensity-specific distributions and intensity-specific variances. We presented a method for estimating these ISDs and ISVs from an image (or a collection of images) directly, and illustrated how the ISVs can be used to define intensity-specific distance measures. Next, we applied these concepts to the aforementioned applications, and compared them to the standard approaches. Results from both quantitative and qualitative analyses showed an increased performance, confirming the functioning and effectiveness of the proposed concepts. Based on these results, together with the method's simplicity, we believe the proposed concepts to be valid and useful tools for (pre)processing heteroscedastic medical images. The applications discussed above can be used on their own or as a preprocessing step, whereas the proposed similarity measure can easily be incorporated into new and existing frameworks.

ACKNOWLEDGMENT

The authors of this work would like to thank F. Tilotta et al. for permitting the use of their head-CT dataset throughout this work.

REFERENCES

[1] R. Wolz, C. Chu, K. Misawa, M. Fujiwara, K. Mori, and D. Rueckert, "Automated abdominal multi-organ segmentation with subject-specific atlas generation," *IEEE Trans. Med. Imag.*, vol. 32, no. 9, pp. 1723–1730, Sep. 2013.

[2] H. R. Roth *et al.*, "DeepOrgan: Multi-level deep convolutional networks for automated pancreas segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2015, pp. 556–564.

[3] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels," School Comput. Commun. Sci., Ecole Polytechn. Fédérale Lausanne, Lausanne, Switzerland, Tech. Rep. 149300, 2010.

[4] G. Zhang, P. Yan, H. Zhao, and X. Zhang, "A contrast enhancement algorithm for low-dose CT images based on local histogram equalization," in *Proc. IEEE 2nd Int. Conf. Bioinf. Biomed. Eng. (ICBBE)*, May 2008, pp. 2462–2465.

[5] Z. Al-Ameen, G. Sulong, A. Rehman, A. Al-Dhelaan, T. Saba, and M. Al-Rodhaan, "An innovative technique for contrast enhancement of computed tomography images using normalized gamma-corrected contrast-limited adaptive histogram equalization," *EURASIP J. Adv. Signal Process.*, vol. 2015, no. 1, p. 32, Dec. 2015, doi: https://doi.org/10.1186/s13634-015-0214-1

[6] T.-L. Tan, K.-S. Sim, and A.-K. Chong, "Contrast enhancement of CT brain images for detection of ischemic stroke," in *Proc. IEEE Int. Conf. Biomed. Eng. (ICoBE)*, Feb. 2012, pp. 385–388.

[7] A. Padma and R. Sukanesh, "Automatic classification and segmentation of brain tumor in CT images using optimal dominant gray level run length texture features," *Int. J. Adv. Comput. Sci. Appl.*, vol. 2, no. 10, pp. 53–59, 2011.

[8] N. M. Noor, N. E. A. Khalid, M. H. Ali, and A. D. A. Numpang, "Fish bone impaction using adaptive histogram equalization (AHE)," in *Proc. IEEE 2nd Int. Conf. Comput. Res. Develop.*, May 2010, pp. 163–167.

[9] M. Goryawala, S. Gulec, R. Bhatt, A. J. McGoron, and M. Adjouadi, "A low-interaction automatic 3D liver segmentation method using computed tomography for selective internal radiation therapy," *BioMed Res. Int.*, vol. 2014, Jul. 2014, Art. no. 198015, doi: https://doi.org/10.1155/2014/198015

[10] K. Wei, B. He, T. Zhang, and X. Shen, "A novel method for segmentation of CT head images," in *Proc. IEEE 1st Int. Conf. Bioinf. Biomed. Eng. (ICBBE)*, Jul. 2007, pp. 717–720.

[11] Q. Hu, G. Qian, A. Aziz, and W. L. Nowinski, "Segmentation of brain from computed tomography head images," in *Proc. 27th Annu. Int. Conf. Eng. Med. Biol. Soc. (IEEE-EMBS)*, Jan. 2006, pp. 3375–3378.

[12] T. H. Lee, M. F. A. Fauzi, and R. Komiya, "Segmentation of CT brain images using unsupervised clusterings," *J. Vis.*, vol. 12, no. 2, pp. 131–138, 2009.

[13] O. Gambino *et al.*, "Automatic volumetric liver segmentation using texture based region growing," in *Proc. Int. Conf. IEEE Complex, Intell. Softw. Intensive Syst. (CISIS)*, Feb. 2010, pp. 146–152.

[14] X. Zang, Y. Wang, J. Yang, and Y. Liu, "A novel method of CT brain images segmentation," in *Proc. IEEE Int. Conf. Med. Image Anal. Clin. Appl. (MIACA)*, Jun. 2010, pp. 109–112.

[15] P. Joris. *GitHub Repository*. Accessed: May 2, 2018. [Online]. Available: http://www.github.com/philip-joris

[16] F. Tilotta *et al.*, "Construction and analysis of a head CT-scan database for craniofacial reconstruction," *Forensic Sci. Int.*, vol. 191, nos. 1–3, p. 112, 2009.

[17] *Multi-Atlas Labeling Beyond the Cranial Vault—Workshop and Challenge*. Accessed: Mar. 9, 2017. [Online]. Available: https://www.synapse.org/#!Synapse:syn3193805/wiki/89480

[18] R. K. S. Kwan, A. C. Evans, and G. B. Pike, "MRI simulation-based evaluation of image-processing and classification methods," *IEEE Trans. Med. Imag.*, vol. 18, no. 11, pp. 1085–1097, Nov. 1999.

[19] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, and A. C. Evans, "BrainWeb: Online interface to a 3D MRI simulated brain database," *NeuroImage*, vol. 5, no. 4, p. S425, 1997.

[20] *Internet Brain Segmentation Repository (IBSR)*. Accessed: Nov. 22, 2016. [Online]. Available: https://www.nitrc.org/projects/ibsr

[21] J. M. Prewitt, "Object enhancement and extraction," *Picture Process. Psychopictorics*, vol. 10, no. 1, pp. 15–19, 1970.

[22] I. Sobel and G. Feldman, "A 3×3 isotropic gradient operator for image processing," *Pattern Classification Scene Anal.*, pp. 271–272, 1973.

[23] K. Bowyer, C. Kranenburg, and S. Dougherty, "Edge detector evaluation using empirical ROC curves," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 1. Jun. 1999, pp. 354–359.

[24] R. R. Sokal, "A statistical method for evaluating systematic relationship," *Univ. Kansas Sci. Bull.*, vol. 38, no. 22, pp. 1409–1438, 1958.

[25] J. A. Hartigan, *Clustering Algorithms*, vol. 209. New York, NY, USA: Wiley, 1975.

[26] *Statistical Parametric Mapping (SPM)*. Accessed: Aug. 23, 2017. [Online]. Available: http://www.fil.ion.ucl.ac.uk/spm/

[27] R. Hummel, "Image enhancement by histogram transformation," *Comput. Graph. Image Process.*, vol. 6, no. 2, pp. 184–195, 1977.

[28] S. M. Pizer *et al.*, "Adaptive histogram equalization and its variations," *Comput. Vis., Graph., Image Process.*, vol. 39, no. 3, pp. 355–368, 1987.

[29] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2006.

[30] M.-L. Antonie, O. R. Zaiane, and A. Coman, "Application of data mining techniques for medical image classification," in *Proc. 2nd Int. Conf. Multimedia Data Mining*, 2001, pp. 94–101.

[31] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Image registration by maximization of combined mutual information and gradient information," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2000, pp. 452–461.

[32] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. G. Hill, and D. J. Hawkes, "A comparison of similarity measures for use in 2-D-3-D medical image registration," *IEEE Trans. Med. Imag.*, vol. 17, no. 4, pp. 586–595, Aug. 1998.

[33] K. Haris, S. N. Efstratiadis, N. Maglaveras, and A. K. Katsaggelos, "Hybrid image segmentation using watersheds and fast region merging," *IEEE Trans. Image Process.*, vol. 7, no. 12, pp. 1684–1699, Dec. 1998.

[34] B. N. Li, C. K. Chui, S. Chang, and S. H. Ong, "Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation," *Comput. Biol. Med.*, vol. 41, no. 1, pp. 1–10, 2011.

[35] L. He *et al.*, "A comparative study of deformable contour methods on medical image segmentation," *Image Vis. Comput.*, vol. 26, no. 2, pp. 141–163, 2008.

[36] T. F. Cootes, A. Hill, C. J. Taylor, and J. Haslam, "Use of active shape models for locating structures in medical images," *Image Vis. Comput.*, vol. 12, no. 6, pp. 355–365, 1994.

[37] E. A. Riskin, T. Lookabaugh, P. A. Chou, and R. M. Gray, "Variable rate vector quantization for medical image compression," *IEEE Trans. Med. Imag.*, vol. 9, no. 3, pp. 290–298, Sep. 1990.

[38] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 187–198, Apr. 1997.

[39] W. M. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Med. Image Anal.*, vol. 1, no. 1, pp. 35–51, 1996.

[40] S. Gupta, A. Agrawal, K. Gopalakrishnan, and P. Narayanan, "Deep learning with limited numerical precision," in *Proc. 32nd Int. Conf. Mach. Learn. (ICML)*, 2015, pp. 1737–1746.

[41] J. C. Bezdek, W. Full, and R. Ehrlich, "FCM: The fuzzy c-means clustering algorithm," *Comput. Geosci.*, vol. 10, nos. 2–3, pp. 191–203, 1984.

[42] K. V. Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "Automated model-based tissue classification of MR images of the brain," *IEEE Trans. Med. Imag.*, vol. 18, no. 10, pp. 897–908, Oct. 1999.

PHILIP JORIS received the bachelor's and master's degrees in applied engineering from Artesis Antwerp, Belgium, in 2012, and the master's degree in artificial intelligence from KU Leuven, Belgium, in 2013, where he is currently pursuing the Ph.D. degree in electrical engineering. He is currently with the Medical Imaging Research Center. His current research is focused on applying image computing techniques in forensic applications, more specifically in the fields of bloodstain pattern analysis and virtual autopsy.

WIM DEVELTER received the degree from KUL in 2015 as a clinical pathologist. He became a Staff Member with the Forensic Department, Catholique University of Leuven (UZ-KUL), in 2008, after his medical school formation. He participated in many missions with the Belgian Disaster Identification Team and represents the Belgian Pathologists, Interpol. In 2010, he achieved the first ISO 17020 accreditation of the forensic autopsy. He is currently the Head of the Forensic Pathology Division and he is also responsible for the fetal and neonatal pathology within the Pathology Department, UZ-KUL. His current research interest includes anthropology, quality control, and postmortem imaging. Since histology is an important part of the medicolegal investigation, he focused on histology and molecular pathology.

WIM VAN DE VOORDE received the M.D. and Ph.D. degrees in medicine from the University of Leuven in 1984 and specialized in anatomical pathology from University Hospitals UZ Leuven and the Ph.D. degree in prostate cancer from the Department of Pathology, UH Leuven, in 1995. He was a Staff Member with the Department of Pathology, UH Leuven, in 1990. From 1995 to 1996, he was a Post-Doctoral Resident with the Institut für Rechtsmedizin, University of Zurich. In 1997, he started up the Department of Forensic Medicine, UH Leuven. He is currently a Coordinating Supervisor in forensic medicine with five recognized internships for doctor-specialists in training in forensic medicine. Since 2006, he has been the Head of the forensic medicine medical service with the University Hospitals UZ Leuven. He was a Full Professor (part-time) with the Faculty of Medicine, University of Leuven, where he is currently the Head of forensic biomedical sciences with the Department of Imaging and Pathology, and also the Chair in forensic medicine/forensic medical sciences with the faculty of Medicine. He also holds the Chair of medical ethics (medical deontology). He teaches at the Belgian National Police Academy and several police schools. He is also a medical specialist in the field of anatomical and forensic pathology, forensic pathologist, a member of the recognition committee for legal medicine, a member of the committee for legal costs in criminal cases, a Former President of the Royal Belgian Society for Forensic Pathology, a member (expert) of the standards and assessment advice committee Forensic Pathology of the Netherlands Register of Court Experts, The Netherlands, a member of the Forensic Medical Society and the International Academy of Legal Medicine, and a Former Expert/Trainer at the EU twinning project Improving the Skills of Forensic Experts, Turkey, from 2011 to 2013.

PAUL SUETENS is currently a Full Professor and the Head of the Division Image and Speech Processing with the Department of Electrical Engineering, KU Leuven. He is also the Chairman of the Medical Imaging Research Center, University Hospital Leuven. His research focuses on medical imaging and medical image computing. He has authored over 500 peer-reviewed papers in international journals and conference proceedings and author of the book *Fundamentals of Medical Imaging*, three editions, in 2002, 2009, and 2017, respectively. Methodologically, this research belongs to the domains of computational science and machine learning. He is involved in several graduates (biomedical engineering and biomedical sciences) and postgraduate (advanced medical imaging and biomedical engineering) programs. He has been a member of several research councils and commissions and has been an associate editor of the IEEE TRANSACTIONS OF MEDICAL IMAGING since 2004. He is currently a Co-Founder of two spin-off companies of the KU Leuven (Medicim in 2002 and icoMetrix in 2011).

FREDERIK MAES received the master's degree in electrical engineering from KU Leuven in 1991, the M.Sc. degree in electrical engineering from Stanford University in 1992, and the Ph.D. degree in electrical engineering from KU Leuven in 1998. He was a Professor with the Division Processing Speech and Images, Department of Electrical Engineering, KU Leuven, in 2004. His main research interest is on computational strategies for medical image analysis and their application for solving challenging research questions in clinical and biomedical research in a broad variety of disciplines, including neuro-, cardiac and oncological imaging, radiotherapy, and preclinical small animal imaging. His research activities are embedded within the Medical Imaging Research Center, University Hospital UZ Leuven, which allows for close interaction with clinical and biomedical researchers, resulting in many formal and informal collaborations and strong leveraging of his research towards clinically relevant applications.

DIRK VANDERMEULEN received the master's degree in computer science and the Ph.D. degree in electrical engineering from KU Leuven, Belgium, in 1983 and 1991, respectively. His initial research concentrated on image processing for computer-assisted stereotactic neurosurgery. Some of his work was implemented in a commercially available stereotactic planning system. His work then moved to medical image analysis, still with a strong focus on neurosurgical and neurological applications. He became an Assistant Professor in 1998 and an Associate Professor in 2004 with the Division for Image and Speech Processing, Department of Electrical Engineering, KU Leuven, where he co-supervises research work on model-based image analysis, with a strong emphasis on image registration and segmentation. Since 2002, he has been involved in applying computer vision and medical image analysis techniques to forensic imaging, more in particular cranio-facial reconstruction, biometric authentication and facial analysis. Part of his work was performed in collaboration with research institutes abroad Telecom Paris, in 1993, INRIA, Sophia Antipolis, in 2000, the University of Melbourne in 2010, and the University of Pretoria, from 2016 to 2017. He is also appointed as an Extra-Ordinary Professor with the University of Pretoria, South Africa.

PETER CLAES received the Ph.D. degree in engineering with a focus on medical image computing from KU Leuven, Belgium, in 2007. During his Ph.D. he developed a computer-based craniofacial reconstruction approach for victim identification purposes. After his Ph.D., he established his own research trajectory and vision, with fundamental interest in pattern recognition and predictive modeling within computational imaging and biology. He held a Postdoctoral position with the Melbourne Dental School, University of Melbourne, Australia, until 2011. During that time, he built an international and versatile network of collaborations that is still very useful and active today. He is currently appointed as an Honorary Fellow with the Murdoch Children's Research Institute, Australia, and as a Research Expert with KU Leuven, ESAT/PSI/MIC, Belgium.