Segmenting Leukomalacia using Textural Information and Mathematical Morphology

Ewout Vansteenkiste[1], Alessandro Ledda[1], Gjenna Stippel[1], Bruno Huysmans[1], Paul Govaert[2], Wilfried Philips[1]
[1]Ghent University, Dept. of Telecommunication and Information Processing
Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium
[2]Sophia Children's hospital EMC, Dept. of Neonatology, Rotterdam, The Netherlands
ewout.vansteenkiste@UGent.be, alessandro.ledda@UGent.be

Abstract— **In this article we present a technique for segmenting the affected tissue visible as white flaring in the ultrasound brain images of neonates with Leukomalacia (White Matter Damage). The technique combines both textural information of the investigated tissue as well as mathematical morphology in order to detect and delineate the boundaries of the affected parts of the brain. The reproducibility of the proposed technique is evaluated and the experimental results are validated by comparing them to the manual delineations of 12 expert medical doctors from different institutions. Although it seems hard to reach a consensus on the correct segmentation of the flaring, because of the lack of a golden standard in the ultrasound images, we show that our method outperforms the existing techniques based on active contours in speed and is more accurate.**

Keywords— **Medical Ultrasound, Leukomalacia, segmentation, mathematical morphology**

I. Introduction

20 To 50 percent of the neonates with a very low birth weight (< 1500 g) suffer from the "White Matter Damage" (WMD) [1]. Due to a lack of oxygen in the brain parts of the white matter die. This leads flaring on the ultrasound image as shown in figure 1. In current practice, experts depend on visual inspection of ultrasonic (US) images for the diagnosis of WMD [1] when there is no other information such as MRI available.

In order to objectively support the diagnosis (e.g. in medical reports, for purposes of meta-analysis or to objectively evaluate the progression of the disease in time, staging) there is a clear need for (semi-) automated algorithms for delineating (segmenting) the affected regions. A classical and commonly used approach to the delineation problem in other imaging modalities is the technique of active contours [2]. However, this technique performs poorly on US images, because of speckle noise. The contours tend to get stuck on isolated speckles [3]. Another problem is that US images acquired using different equipment can differ substantially which implies that algorithms have to be tuned, often even redesigned for different equipment.

In this paper, we present a new technique that is more robust against speckle noise and that is machinery independent (it can be used on different types of ultrasound machines without refining). The segmentation technique is based on the experimental observation [3] that textural features based on the co-occurrence matrix can to some degree distinguish between (large) affected and non affected areas. Some morphological image processing is used to detect the outer boundaries of the textural homogeneous affected parts.

To validate our new technique, we performed segmentation tests on a population of 21 images from WMD patients and 7 images from patients without this disease. The images were acquired using an Ultramark ultrasound machine. To validate the machine-independence, we also tested the algorithm on 35 images acquired using an Acuson ultrasound machine. The results show that better delineations are achieved with this technique than with the active contours, and that it is not necessary to adapt or tune the algorithm on different equipment.

Our algorithm is fast which could be useful in case of real-time post processing applications like 3D ultrasound segmentations or 3D volumetrical measurements of the flares.

In the next sections we will consequently explain the method used and the segmentation algorithm before turning to the results, the validation of those and their discussion. We will end with the further research that remains to be done.

II. Method

A. Textural information

According to earlier results [3] and [4] textural features based on the co-occurrence matrix are useful to distinguish affected from unaffected brain tissue. The co-occurrence matrix counts the number of times $n(a, b)$ two pixels $p(x, y) = n_1$ and $p(x + a, y + b) = n_2$ located in a fixed relative position to each other assume a particular combination (n_1, n_2) of grey values. As such it is a two-

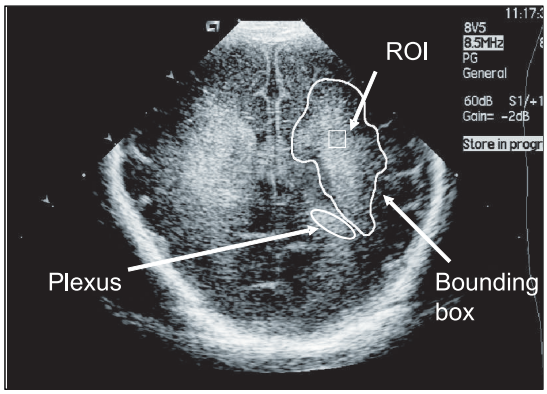

Fig. 1. Brain infected with WMD, one affected part, delineated white flare, is shown in a bounding box. A square region of interest (ROI) is also shown for texture examination. The Plexus Chorideus is also delineated

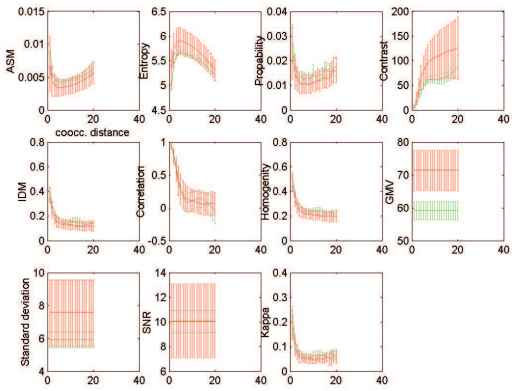

Fig. 2. Curves of co-occurrence parameters over multiple distances, with error bars. Darker curve (red if color) is for the affected images, lighter one (if color green) for healthy ones.

dimensional histogram [5]. In practice, the co-occurrence matrix is computed in a square region, one or more features (numbers) are derived from it and these features are used to classify the region. In earlier work usually two particular features, the mean grey value and the contrast were used [3].

However, many other combinations of features are possible. In order to investigate the best combination of features for our purpose, we have evaluated 10 first and second order features (Angular Second Moment, Entropy, Contrast, Correlation, Inverse Difference Moment, Max. Probability, Mean Grey Value, Standard Deviation, Signal-to-noise ratio and Kappa [5]) derived from the co-occurrence matrix to investigate if adding these could features would lead to a better classification. We repeated this evaluation for different values of $a = d\cos\theta$ and $b = d\sin\theta$, i.e. for $\theta = 0, 45, 90$ and 135 degrees and for d in the range 1 to 20 for window sizes (ROI) ranging from 30x30 up to 60x60 pixels. These regions in which the co-occurrence matrix is computed, were selected in the proper anatomical region where WMD usually occurs.

Figure 2 shows us some results from this experiment, in which we see the average value (with the standard deviation as vertical bars) of the features over the unaffected (green) and affected (red) images for different distances in the co-occurrence matrix for a specific angle (0 degrees) and window size (30x30 pixels). The figure shows that for any d only the grey mean value is an acceptable discriminating feature, at least in the case where only one parameter is used, combinations of parameters are not discussed here [3]. Similar results were obtained with the other window sizes and angles. The results imply that the structural texture pattern is not very different in affected and unaffected tissue at least when we use these features, that adding extra texture features does not drastically improve the classification and that it is best to rely mainly on the average grey value for classification.

Given the rather poor textural results we applied another technique based on prior research involving the correlation in grey mean value between the flaring and the plexus chorideus, an anatomical feature partly inside the ventrical as can also be seen in figure 1. In [6] it is shown that there is a correlation between the grey mean value of the plexus and of the flaring around.

B. Segmentation algorithm

The information above can be used in the development of a semi-interactive algorithm that requires the user to indicate a rough bounding box around the flare of interest after which the algorithm computes the exact boundary. This type of interaction is smaller compared to the interaction required in the active contours technique where typically multiple points have to be precisely chosen.

More into detail the technique proceeds as follows: first of all, the user selects 2 Regions of Interest or ROI, figure 1. The first ROI contains the plexus chorideus which is easy to segment and the second ROI contains the desired flaring. Now since the borders of the flaring are unclear and a manual segmentation is rather inaccurate, some background is usually included [3] in the second ROI.

The first ROI can be seen as a bounding box containing the reference grey mean value, the second as the area where the flaring is expected. Next, a threshold will is applied. All pixels in the flaring with grey value less than the mean grey value of the plexus ROI are set to zero. This we do based on the results in [6] Segmentation in our case is region based and comes down to classifying a pixel in the class representing the flare or in the class representing background. Therefore, the remaining grey values are

made binary. The upper left part of figure 3 shows the result of this combined pre processing.
The resulting binary image however contains a lot of holes due to the speckled nature of the image. These holes are part of the flaring, but if we make a segmentation on this current image, several little regions will be detected, although we presume the flare is one large connected area. Further image processing is then done with the help of morphological operators.
Mathematical morphology image processing is based on set theory, so the shapes of objects in a binary image are represented by object membership sets. Morphological operations can simplify image data, preserving the objects' essential shape characteristics, and can eliminate irrelevant objects. Our technique is based on two basic widely used operations: a dilation, which fills holes and smooths out the contour lines, and an erosion, which removes small objects and disconnects objects connected by a small bridge. Such operations are defined in terms of a structuring element, a small window that scans the image and alters the pixels in function of its window content. The choice of the proper structuring element is important, most commonly little rectangles, squares, discs or crosses are used depending on the application and morphological properties of the images. A dilation of image A with structuring element B ($A \oplus B$) blows up the object, an erosion ($A \ominus B$) lets it shrink. Dilation and erosion are union and intersection operations, respectively:

$$A \oplus B = max\{A(a-b) + B(b) \mid b \in B,\ a-b \in A\}$$
$$A \ominus B = min\{A(a+b) - B(b) \mid b \in B, a+b \in (A)\}$$

where a is each pixel of the image A and b is each pixel in the structure element B. Other operations, like the *opening* $A \circ B$ (an erosion followed by a dilation) and the *closing* $A \bullet B$ (a dilation followed by an erosion), are derived from the basic operators. The closing, in specific, acts like a dilation (it fills holes), but does not blow up the objects, so the original object size is retained.
How we will use these operators in our technique will now be explained. After the preprocessing we still notice tiny black holes and outlying speckles. Using a square or a disk as structure element we can fill those holes and connect the speckles by closing the image with these structure element as can also be seen in figure 3, upper right en lower left. Since there is no dominant direction in the flaring using disks gives a more natural result, so from now on stick to them. After the closing we can already measure the pixel area of the flare, extracting the contour of the flare can be done with another morphological operation, the gradient. Taking the difference of the dilation with the erosion, thus calculating $G^B(A) = (A \oplus B) \backslash (A \ominus C)$ gives us this

Fig. 3. Upper left: preprocessing result, Upper right: closure with a disk , Lower left: closure with a squared structure element, Lower right: gradient operation.

gradient. In this equation A is the upper right image in 3, B is a disk of radius 4 to 6 for the dilation and C is a disk of radius 1 to 3 for the erosion. The result of the gradient operation can be seen in the lower right part of figure 3. This will be our final delineation. It is possible that for different images different radii of the structure elements have to be tested in order to obtain best result. The interval mentioned above (4 to 6, and 1 to 3 are good indications in this).

III. Experimental Results

When compared to a standard active contour method [2], our method outperforms it as well in speed as in the results obtained. As can be seen from figure 5 and figure 4, our method is less sensitive to isolated speckles surrounding the flares which makes the contour more refined. Our method can also be used over different machinery without adaption, as for now we tested it on images obtained with Ultramark and Acuson equipment. In term of time our algorithm needs about 1 sec. where the active contours go up to 20 sec. (AMD, 2GHz. processor).

IV. Validation

A. Inter-observer variability

In order to compare our method to the manual delineations made by experts we set up the following experiment. We collected a set of 10 images of children infected by WMD and sent them to 12 medical doctors from differ-

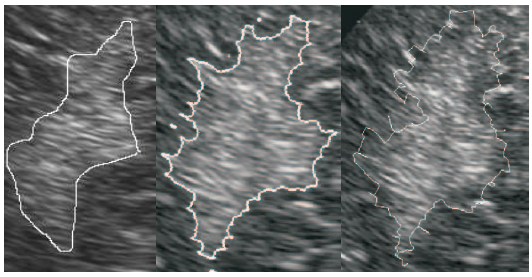

Fig. 4. On the left the delineation of a medical doctor, in the middle our technique, on the right the active contour method performed on an Ultramark ultrasound machine

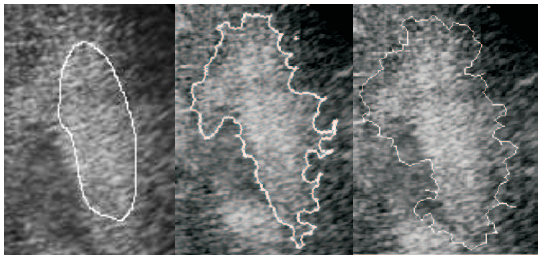

Fig. 5. On the left the delineation of a medical doctor, in the middle our technique, on the right the active contour method performed on an Acuson ultrasound machine

Fig. 6. The upper image is one of the 10 used in the test set, in the lower image the different delineations are shown using different colors, each medical doctor has his own color against a blue background

	pix. area method	pix. area MD	*OR* (%)
image 1	10785	6645	60,17
image 3	10785	9794	77,61
image 4	11793	11138	74,90
image 5	12817	12457	83,74
image 6	11792	10890	79,52
image 7	6875	7430	85,22
image 8	6697	5689	69,23

TABLE I
INTER-OBSERVER VARIABILITY

ent institutions, asking them to segment the flaring as they see them, without further prior information. The results were somehow surprising as can be seen in figure 6. This figure shows the overlay of the different delineations of the left and right flaring in the image placed above. As can easily be seen, the inter-observer variability is enormous. The mayor explanation for this is the lack of a golden standard for characterizing affected tissue in ultrasound images, Another reason is the lack of prior medical information, as we said we asked the medical doctors involver to segment blindly. This is a drawback for the validation of the technique.

As for now, we compared our segmentations to those of the medical doctor who acquired the images, as we may expect that he has the best knowledge of the pathology in these images, having access to other complementary medical information of the patients.

In order to do so we calculated the ratio of the area of overlap of different segmentations and the total area of the union of both segmentations. Let A be the manual delineation by the medical expert and B be our delineation, we computed following simple overlap ratio (OR)

$$OR = \frac{A \cap B}{A \cup B}. \tag{1}$$

The results for 8 of the 10 test images can be found in table I, in which we compared the delineation of the medical doctor to the delineation make by our algorithm. The higher the ratio, the better the segmentations overlap. Note that in fact next to the overlap we also have to take the morphology of the segmentations into account. More sophisticated methods of comparing the segmentations, based on [7], which also take into account the morphology are currently investigated.

B. *Intra-observer variability*

In order to measure the intra-observer variability or, i.e. the reproducibility of our method, we set up the following

	pix. area method	pix. area MD	*OR* (%)
image 1	18739	5.1	85,21
image 2	16134	4.0	86,16
image 3	22573	4.2	87,45
image 4	22452	5.2	85,45
image 5	13179	5.9	84,13
image 6	14988	6.5	86,52
image 7	14659	6.8	79,12
image 8	22435	3.9	85,78

TABLE II
INTRA-OBSERVER VARIABILITY

experiment. Out of the test set of 10 pictures, we let the computer pick an image which was presented to the same user to segment. This continued until each of the 10 pictures was picked and segmented 5 times. We did this in order to scramble the images and prevent the user of using prior information if he or she would be asked to segment the same image multiple times. In other words if he or she would segment the same images 5 times the one immediately after the other the chance of having highly correlated segmentations is big. We try to overcome this by presenting the images somehow randomly. The results for the right flaring can be found in table II for 7 of the 10 images. Further research of how the morphology of the segmentations behaves is also here in process. This can partly be done using the overlap region method as explained earlier, only now we take the intersection and union of the 5 segmentations per image. These figures are also presented in II.

V. DISCUSSION

As shown above we have developed an easy-to-use, quite fast, semi-automatic tool to aid medical doctors in their diagnoses of WMD. The direct usefulness of the technique should be seen as mentioned in the introduction in terms of staging information of the disease other than in the precise hard classification and segmentation of the tissue. The pathology changes over time as so the flaring changes.

When we now turn to tables we can conclude from table II that concerning the intra-observer variability, the area of the segmented flares does not differ more than 5 percent in most cases, which is acceptable [7]. Looking at the overlap ratios we might tend to say these are not as high as one would expect, all lay around 85 percent with the exception of one outlier, there where in fact they all do not differ more than 5 percent from the mean. More advanced techniques of comparing the morphology of the delineations could give us extra information on this point.

When we turn to the inter-observer variability and the results in table I, there we obtain Overlap Ratio numbers in the 60 percent (which is rather bad) to 85 percent (which is acceptable) interval. The differences here can be explained by the fact that some medical doctors only segment the core of the flaring, there where other times they also include the outlying, may be less affected tissue. Our technique always tries to find the outliers, unless the bounding box is especially chosen else wise. In the images 1, 4 and 8 of our test set this difference is clearly visible in the Overlap Ratio.

VI. FURTHER RESEARCH

Given the results there remain considerable problems to be tackled, both from the medical as well as the image processing point of view. Medically, the most important drawback is the lack of a golden standard in segmenting the flares in ultrasound images. As can easily be seen from figure 6 the inter-observer variability is very large, this makes it hard to validate the techniques since there is little agreement on what is the standard. All medical doctors segment the core of the affected tissue, in the upper part, right above the ventricles, see figure 6. Only few also segment some outlying tissue also as being affected. The only ways to overcome this problem are or to only look at the core part and try to classify the tissue there, to look for a segmentation band instead of a boundary, or to make a multi-modal registration for instance using MRI images. These drawbacks do influence the image processing problems also, where we still investigate whether texture is a good discriminant for WMD. As may be clear we have not yet found very good distinctive textural features. The nature of the ultrasound images, mostly the speckle included makes this far from easy. In order to improve classification as well as segmentation we investigate preprocessing, before calculation, as well as other textural features based on Gabor Filters and Fractal dimensions.

Another research topic in the development of better techniques for comparing segmentations [8]. This might eventually also help in our search for a golden standard for segmentations.

VII. ACKNOWLEDGEMENTS

We would like to thank all medical doctors who assisted in segmenting the test set and without whom this research would not be possible, M.D. Paul Govaert (Dept. of Neonatology, Sophia Children‘s hospital Rotterdam), M.D. Kris De Coen and M.D. Alexan-

dra Zecic (Dept. of Neonatology, University Hospital, Ghent), Prof. Linda De Vries and M.D. Floris Groenendael (Dept. of Neonatology, Wilhelmina Children‘s Hospital/University Medical Center, Utrecht), M.D. , Ph.D Gerda van Wezel-Meijler,(Dept. of Pediatrics, Division of Neonatology, University Medical Center, Leiden), M.D. K.D. Liem and M.D. R.A. Mullaart (University Children‘s Hospital, Nijmegen), M.D. Frances Cowan (Dept. of Pediatrics and Neonatal Medicine, Imperial College School of Medicine, Hammersmith Hospital, London), Prof. Terrie Inder (Royal Women‘s and Children‘s Hospital, Melbourne), M.D. Andrea Mewes and M.D. Goetz Welsch (Bringham and Women‘s Hospital, Boston).

REFERENCES

[1] J.J. Volpe “Neurologic outcome of prematurity“ *Arch. Neurol.*, pp 297-300, 1998.

[2] C. Zu and J.L. Prince, “Gradient vector flow: a new external force for snakes,” *Conference on computer Vision and Pattern Recognition (CVPR ‘97)*, pp. 66–71, 1997.

[3] G. Stippel, W. Philips, A. Zecic, P. Govaert, I. Lemahieu, I. Duskunovic, “A new filtering method for ultrasound images incorporating prior statistics concerning medical features,” *Proceedings of ICIP 2001*, 2001.

[4] Rotteveel Thijssen, Geemen and Mullaart., “Characterization of echographic image texture by co-occurrence matrix parameters,” *Ultrasound in Med. and Biol.)*, vol. 23, pp. 559–571, 1997.

[5] K. Shanmugan R.M. Haralick and I. Dinstein, “Textural features for image classification,” *IEEE Trans. Systems, Manufact., Cybernet.*, vol. 3, pp. 610–621, 1988.

[6] G. Stippel, I. Duskunovic, W. Philips, I. Lemahieu, P. Govaert, A. Zecic “Segmenting ’Flares’ in ultrasound images using prior statistics” *Image Processing and Communication.)*, vol. 7, nr. 1-2 pp. 41-54, 2000.

[7] V. Chalana, Y. Ki, “A Methodology for Evaluation of Boundary Detection Algorithms on Medical Images,” *IEEE Trans. on Medical Imaging*, vol. 16, nr. 5 pp. 642–652, 1997.

[8] S.K. Warfield, K.H. Zou, M.R. Kaus, W.M. Wells “Simultaneous validation of image segmentations and assessment of expert quality” *Proc. of International Symposium of Biomedical Imaging: Macro to Nano*, Washington DC, USA 2000.