Three dimensional multi–scale visual words for texture–based cerebellum segmentation

Antonio Foncubierta–Rodríguez[a], Adrien Depeursinge[a,b], Laura Gui[b,c] Henning Müller[a,b]

[a]University of Applied Sciences Western Switzerland (HES–SO), Sierre, Switzerland;
[b]Department of Radiology and Medical Informatics, University of Geneva and Geneva University Hospitals (HUG), Geneva, Switzerland;
[c]Faculty of Medicine, University of Geneva, Geneva, Switzerland.

ABSTRACT

Segmentation of the various parts of the brain is a challenging area in medical imaging and it is a prerequisite for many image analysis tasks useful for clinical research. Advances have been made in generating brain image templates that can be registered to automatically segment regions of interest in the human brain. However, these methods may fail with some subjects if there is a significant shape distortion or difference from the proposed models. This is also the case of newborns, where the developing brain strongly differs from adult magnetic resonance imaging (MRI) templates.

In this article, a texture–based cerebellum segmentation method is described. The algorithm presented does not use any prior spatial knowledge to segment the MRI images. Instead, the system learns the texture features by means of a multi–scale filtering and visual words feature aggregation. Visual words are a commonly used technique in image retrieval. Instead of using visual features directly, the features of specific regions are modeled (clustered) into groups of discriminative features. This means that the final feature space can be reduced in size and also that the visual words in local regions are really discriminative for the given data set. The system is currently trained and tested with a dataset of 18 adult brain MRIs. An extension to the use with newborn brain images is being foreseen as this could highlight the advantages of the proposed technique.

Results show that the use of texture features can be valuable for the task described and can lead to good results. The use of visual words can potentially improve robustness of existing shape–based techniques for cases with significant shape distortion or other differences from the models. As the visual words based techniques are not assuming any prior knowledge such techniques could be used for other types of segmentations as well using a large variety of basic visual features.

Keywords: texture–based segmentation, texture features, visual words, multi–scale image processing

1. INTRODUCTION

Medical imaging has evolved strongly over the past 30 years and images are used to support decisions in not only in radiology.[1,2] To aid also less experienced physicians to image analysis–based decision support is an important research domain. For many applications the regions of interest in medical images are very small and structures in the images need to identified to concentrate the processing onto these regions of interest.[3] Segmentation is thus often a first preprocessing step in medical image analysis to concentrate the further processing onto the areas of interest.[4]

Thus, segmentation is often regarded as one of the major preprocessing tasks in medical imaging. It allows further study of precise regions of interest for a large variety of clinical applications. Quantitative analysis of the various structures present in the brain can lead to better diagnosis or treatment planning in diseases such as for Alzheimer[5] or brain tumors. The techniques described in this article focus on the analysis of the cerebellum region in brain MRI (Magnetic Resonance Imaging), which has been an intensive research area over the past years and is known to have relation to diseases such as age–related degeneration in adults[6] or supratentorial brain injury in newborns.[7]

Further author information: (Send correspondence to Antonio Foncubierta, Email: antonio.foncubierta@hevs.ch)

Manual segmentation of brain structures, while producing excellent results based on the experience of the user, is often time–consuming and thus does not allow to work on large–scale data sets. In contrast, automatic segmentation can lead to relatively accurate results and is much faster, thus allowing to treat much larger amounts of data and compare the results across many cases. Many of the automatic segmentation techniques for MRI brain images are partly or totally based on atlases, templates or other types of spatial priors built upon statistical information from several subjects.[8,9] Such model–based segmentation can improve results strongly when cases are relatively standard but has a risk to deliver poor results in situations where the model building is difficult.

The cerebellum structure in the brain, the main target of this article, consists of two regions: cortex and white matter. Texture plays an important role in the visual aspect of these regions. Whereas the white matter has a light, homogeneous appearance, the cortex contains rich, dark patterns. Figure 1 shows an example of these two regions.

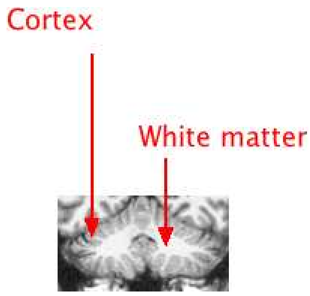


Figure 1. Example slice showing the visual appearance of cerebellum cortex and white matter.

The state of the art for cerebellum segmentation often neglects this texture information and relies only on prior knowledge such as probability maps. Even methods specifically designed to take advantage of texture to segment the cerebellum, rely most often on atlases[10] or user interaction.[11] Problems arise when there are insufficient amounts of data to create an atlas or when the shape and spatial appearance of the subjects are largely variable.

In this text a segmentation method is presented that makes no use of any prior spatial knowledge for the segmentation. The system learns the three–dimensional texture patterns present in the cerebellar region from a subset of training patients and classifies the voxels of new images based on the texture features. The method is evaluated with a freely–available adult brain MRI dataset. The currently used data set is small and having more training samples can potentially increase the performance of the segmentation in an important way.

One of the future goals of the article is the integration of the segmentation and local analysis for image retrieval.[3,12] Retrieval in the case of brain images or the cerebellum needs to concentrate on small local areas to detect abnormal regions and has to use 3D information[13,14] for the retrieval of the cases and include clinical data for case–based information retrieval[15–17] in an environment using information fusion.[18]

This article is organized as follows: Section 2 explains the data sets used and the main techniques that are employed in the article, Section 3 lists the main results that are then discussed in Section 4, before the article concludes with Section 5

2. MATERIALS AND METHODS

In this section, information on the image database and the theories supporting the experiments are presented. The framework of the system is based on two main ideas: the 3D wavelet transform for multi–scale texture feature extraction and visual words[19,20] as descriptors, which are based on the patterns actually occurring in a particular data set.

2.1 Adult brain data set

An MRI database of brain images of 18 adult subjects* is used in this article. For each subject, T1–weighted volumetric images that are positionally normalized into the Talairach orientation (rotation only) are available. Image resolutions vary between $\delta_x, \delta_y = 0.837$ to 1.0 millimeters per pixel, and $\delta_z = 1.5$ millimeters in all cases. An expert segmentation is available for each subject. For evaluation, cerebellum cortex and cerebellum white matter are considered and no distinction between left and right hemispheres is made. An example of the images contained in this dataset and their annotated ground truth can be seen in Figure 2.

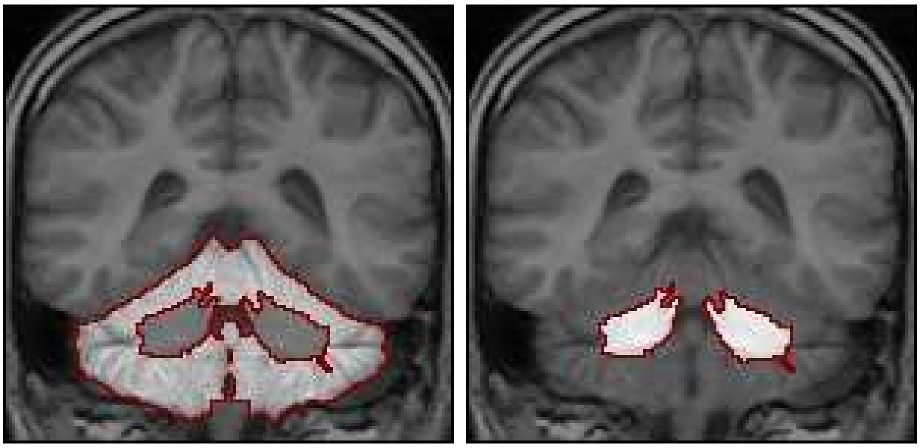


Figure 2. Example slice from the database with annotated cerebellum regions.

2.2 Multi–resolution analysis

Wavelet theory consists of a framework allowing for multi–resolution and multi–scale analysis of images by using scaled and translated versions of a function ψ, the mother wavelet, as shown in Equation 1. A common formulation for the scale definition is based on the minimum number of scales that guarantees that the transform is invertible. In this formulation, scales change according to a dyadic scheme, where $s = 2^j$ defines the current scale.

$$\psi_{s,\boldsymbol{\tau}}(\boldsymbol{x}) = \frac{1}{\sqrt{s}}\psi\left(\frac{\boldsymbol{x}-\boldsymbol{\tau}}{s}\right). \tag{1}$$

Since the image sampling is not isotropic, the multi–resolution analysis is based on the Gaussian function g calculated in physical dimensions by scaling the variables x,y and z using the corresponding values for the voxel spacing in each direction (δ_x,δ_y,δ_z):

$$g_{\boldsymbol{\sigma}}(\boldsymbol{x}) = \frac{1}{\sigma_x\sigma_y\sigma_z\sqrt{(2\pi)^3}}e^{\left(\frac{(x\delta_x)^2}{2\sigma_x^2}+\frac{(y\delta_y)^2}{2\sigma_y^2}+\frac{(z\delta_z)^2}{2\sigma_z^2}\right)}. \tag{2}$$

$$\begin{aligned}\psi(\boldsymbol{x}) &= g_{\boldsymbol{\sigma_1}}(\boldsymbol{x}) - g_{\boldsymbol{\sigma_2}}(\boldsymbol{x}). && (3)\\ \boldsymbol{\sigma_2} &= 1.6\boldsymbol{\sigma_1} && (4)\\ \boldsymbol{\sigma_1} &= \boldsymbol{\sigma_0}2^j && (5)\end{aligned}$$

The choice of using Gaussian functions is made based on their good isotropic properties, which allows to analyze images without making prior choices of orientation. The extracted coefficients are obtained by using the difference of Gaussians (DoG), which provides a good approximation to the Laplacian of Gaussians or Mexican Hat wavelets when the variance parameters $\sigma_{1,2}$ of two Gaussian functions $g_{1,2}$ satisfy $\sigma_2 \approx 1.6\sigma_1$. Since the functions are

*The MR brain data sets and their manual segmentations were provided by the Center for Morphometric Analysis at Massachusetts General Hospital and are available at `http://www.cma.mgh.harvard.edu/ibsr/`.

calculated in physical dimensions, there is no need for having anisotropic variance parameters of $g_{1,2}$. The resulting wavelets are shown in Equation 3. The number of scales used for the wavelet transform is three, with j ranging from 0 to 2. The initial variance parameter is $\sigma_0 = 0.5$ to assure the finest resolution for the wavelet detector at the initial scale.

With each wavelet function, a volumetric image is constructed with the same resolution as the MR image, which showed enhanced texture characterization when compared to decimated wavelet transforms in.[21]

2.3 Visual Words

The term *texture* often has a fuzzy definition and refers to the (sometimes regular or periodic) visual characteristics of the pixel values within a certain region and their relationships, which is not always made explicit by human observers. Since the wavelet transform can describe the transient of the values in the voxel surroundings, a way of aggregating this information for a region of interest is needed. Visual words[19,22] have been widely used in image retrieval and image classification for describing images (or regions of interest) similar to the bag–of–words approach used for text retrieval or text similarity matching.[23] For each voxel, this technique maps a set of continuous low–level features in an area of interest around the pixel, e. g. gray values or wavelet coefficients, into a compact discrete representation consisting of visual words as shown in Figure 3. The visual words are clusters centers in the feature space spanned by wavelet coefficients of similar patterns that actually occur in the data set. This guarantees to have a set of visual features actually corresponding to discriminative patters that do occur in the database. Every image or region is described by the histogram of occurrences of the multi–scale visual–words in a neighborhood around every voxel of this region.

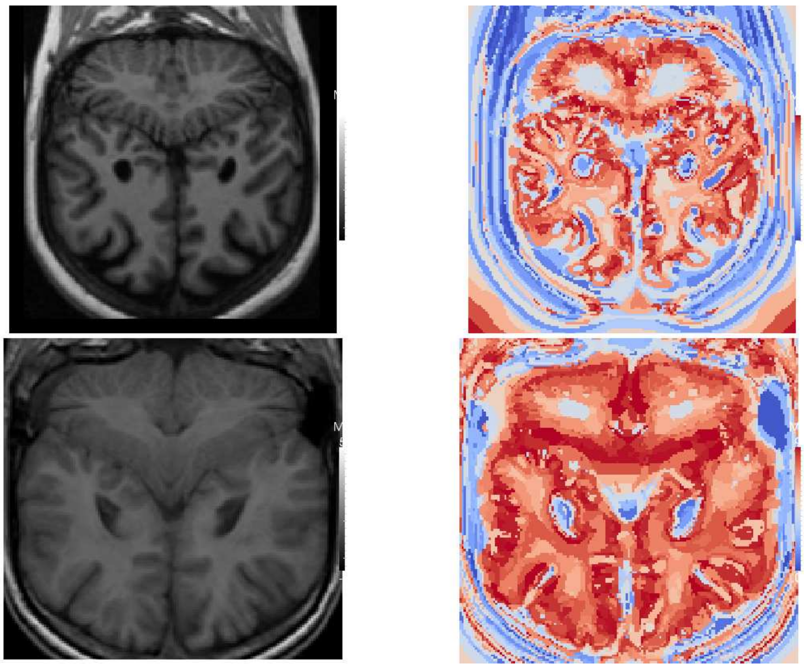


Figure 3. MRI slices and their corresponding representation in the visual words domain.

In this way the local visual word representation maps a feature space for a pixel or region into a space of visually meaningful cluster centers created based on the entire database.

2.4 Gray level histograms

In order to evaluate the suitability of the three–dimensional multi–scale visual words for segmenting the cerebellum, the use of gray level histograms was investigated as well. Gray level histograms describe patterns by using the gray values of a neighborhood around a given voxel, regardless of their relative position. This method has been used in texture description literature[24] and features derived from it proved to be valid for classifying medical images[25] and for many other medical imaging tasks.

2.5 Experimental configuration

The performance of the segmentation is measured by cross–validation separating the data set into training and test data. The classification scheme, shown in Figure 4, is as follows: first, the images are pre–processed in order to reduce the inter–image variability of gray level values among similar regions. This normalization step is achieved by using a histogram equalization filter. Then, the images from the training set are used to compute a visual vocabulary. This is done by clustering the multi–scale feature space computed around each pixel in the image using k–means with $k = 50$ and assigning the 50 cluster centers as the visual words. The value 50 was chosen experimentally and a fairly wide range in terms of size of the visual vocabulary are in principle possible. This visual vocabulary of 50 visual words is used to describe both training and test sets in the visual word domain, by assigning to every voxel the label of the visual word nearest to its multi–scale feature vector. Once the images are described in the visual word domain, a 1–Nearest–Neighbor classifier is trained to classify the voxels into three classes: cerebellum cortex, cerebellum white matter and non cerebellum. The classifier uses the histogram of occurrences of visual words in a cubic block centered in the voxel to be classified. The experiment is repeated with various block sizes.

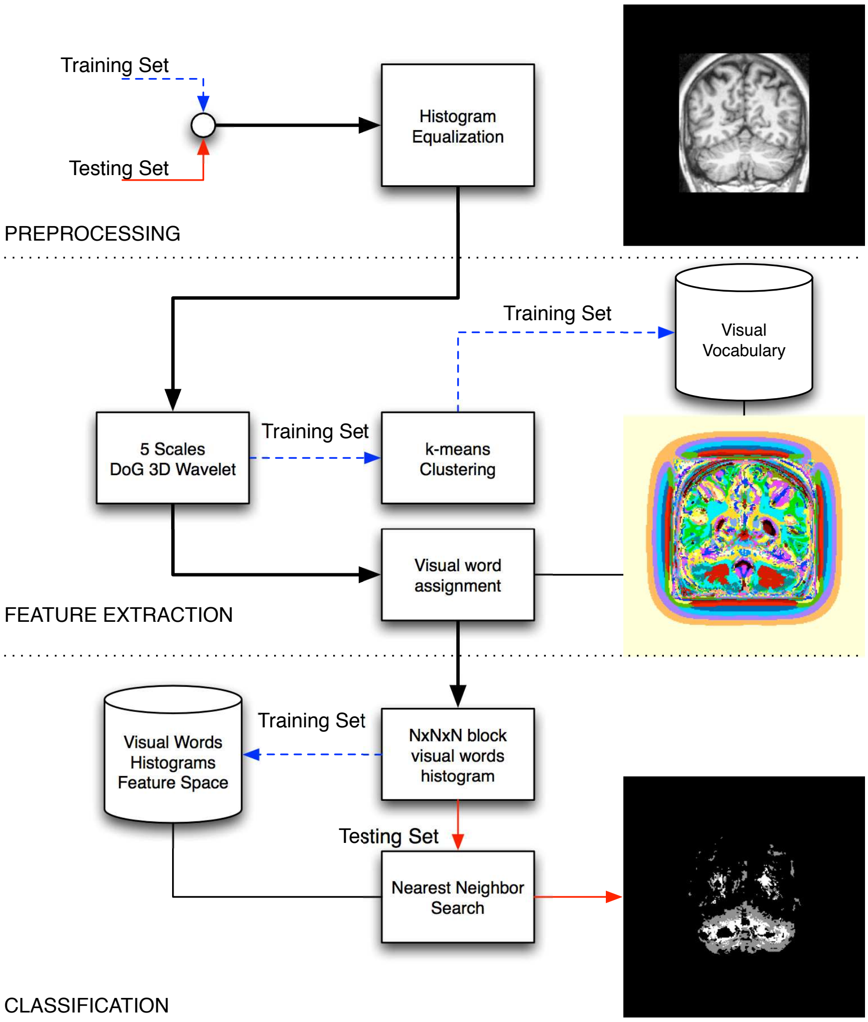


Figure 4. Classification flow diagram using visual words.

For comparison, a similar system based on gray level histograms is also evaluated as shown in Figure 5. In this case, there is no multi–scale feature extraction or visual vocabulary definition. Instead, the images are directly used for classification using the cubic block histogram of gray levels. The number of bins for the histograms is equal to the number of visual words used in the previous experiment. This is done in order to have an equal number of features in the classification step. The histogram analysis serves as a baseline for the evaluation of the system.

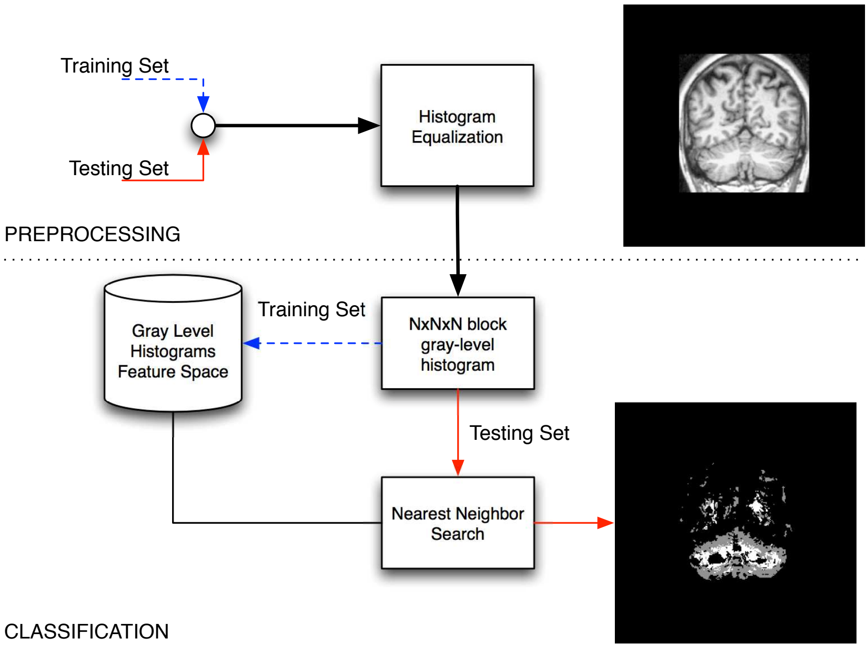


Figure 5. Classification flow diagram using the gray level histogram.

3. RESULTS

This section summarizes the main results of the experiments.

3.1 Performance of the three–dimensional multi–scale visual word configuration

Results for the segmentation were evaluated using four different block sizes for the histogram of visual words, ranging from 5 to 11 pixels. Table 1 shows the confusion matrices for each of the block sizes. Particularly cerebellum cortex and non cerebellum pixels are mixed up. A larger block size leads to better performance for non cerebellum and cortex areas whereas performance for white matter was best with a block size of $7 \times 7 \times 7$. The best performing techniques have approximately 70% of the pixels in each of the classes correctly classified, showing a reasonably good performance.

Table 2 summarizes the results in terms of sensitivity and specificity for each of the classes considered. The overall results show to increase with a larger block size and only for the white matter the specificity is better for a smaller block size than $11 \times 11 \times 11$. Performance gains are in the range of 8–10%, so important differences. Specificity for the best performance with a block size 11 is between 80–90% and sensitivity around 70% showing again the possibility to segment these structures without using any prior models or assumptions.

Figure 6 shows the differences in accuracy as the size of the block increases. The figure underlines the increase of classification accuracy with a varying size of the blocks.

3.2 Performance of the gray level histogram configuration

As a baseline for results comparison a simple histogram intersection using grey level values of the same region sizes and with the same quantization (50 parameters) are used. Results for the gray level histogram configuration

Table 1. Confusion matrices for the multi–scale visual words configuration with varying block size. Labels assigned by the system are shown in columns

(a) $5 \times 5 \times 5$ block

	Non Cerebellum	Cerebellum Cortex	Cerebellum White Matter
Non Cerebellum	**57.44%**	26.77%	15.79%
Cerebellum Cortex	27.86%	**56.41%**	15.73%
Cerebellum White Matter	17.29%	14.94%	**67.77%**

(b) $7 \times 7 \times 7$ block

	Non Cerebellum	Cerebellum Cortex	Cerebellum White Matter
Non Cerebellum	**64.31%**	23.72%	11.98%
Cerebellum Cortex	24.91%	**61.77%**	13.32%
Cerebellum White Matter	14.30%	16.47%	**69.23%**

(c) $9 \times 9 \times 9$ block

	Non Cerebellum	Cerebellum Cortex	Cerebellum White Matter
Non Cerebellum	**71.76%**	19.23%	9.01%
Cerebellum Cortex	24.20%	**62.66%**	13.14%
Cerebellum White Matter	13.66%	17.81%	**68.54%**

(d) $11 \times 11 \times 11$ block

	Non Cerebellum	Cerebellum Cortex	Cerebellum White Matter
Non Cerebellum	**77.34%**	16.05%	6.61%
Cerebellum Cortex	23.89%	**64.09%**	12.02%
Cerebellum White Matter	13.76%	19.88%	**66.36%**

Table 2. Specificity and sensitivity results for each class with the multi–scale visual words configuration

(a) Not cerebellum

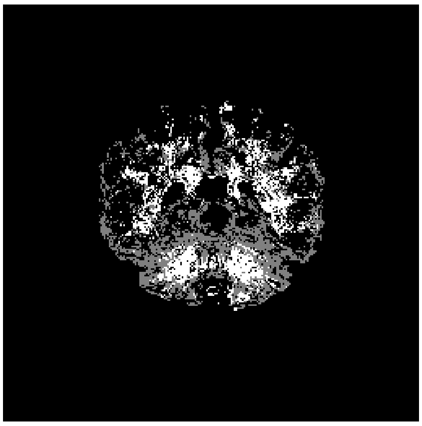
Block Size	Specificity	Sensitivity
5x5x5	73.16%	57.44%
7x7x7	76.12%	64.31%
9x9x9	76.82%	71.76%
11x11x11	**77.09%**	**77.34%**

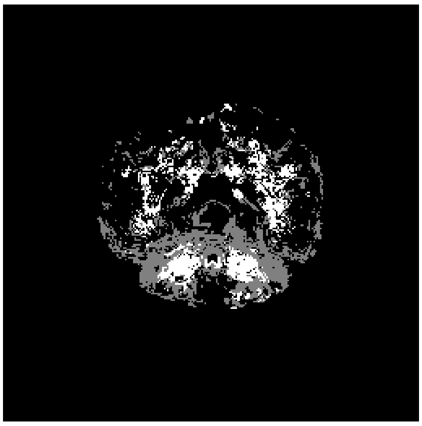
(b) Cerebellum cortex

Block Size	Specificity	Sensitivity
5x5x5	73.35%	56.41%
7x7x7	76.36%	61.77%
9x9x9	80.78%	62.66%
11x11x11	**83.91%**	**64.09%**

(c) Cerebellum white matter

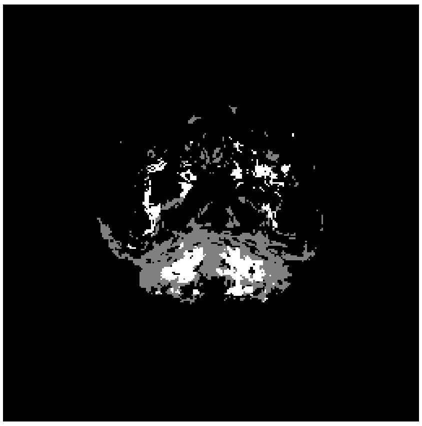
Block Size	Specificity	Sensitivity
5x5x5	84.22%	67.77%
7x7x7	87.91%	**69.23%**
9x9x9	90.62%	68.54%
11x11x11	**92.90%**	66.36%





(a) $5 \times 5 \times 5$ block

(b) $7 \times 7 \times 7$ block



(c) $9 \times 9 \times 9$ block

(d) $11 \times 11 \times 11$ block

Figure 6. Example results of the classification using multi–scale visual words for the same slice and varying block size

are computed for two block sizes, $7 \times 7 \times 7$ and $11 \times 11 \times 11$, only as these obtained the best results for the visual words. Table 3 shows the confusion matrices for these two experimental configurations. The results show that the grey level values are not sufficient to separate particularly white matter and non cerebellum areas of the images. Most of the white matter was incorrectly classified for both block sizes, even worse for the block size of 11.

Table 3. Confusion matrices for the gray level histogram configuration with varying block size. Labels assigned by the system are shown in columns

(a) $7 \times 7 \times 7$ block

	Not Cerebellum	Cerebellum Cortex	Cerebellum White Matter
Not Cerebellum	**61.31%**	24.19%	14.51%
Cerebellum Cortex	36.84%	**54.54%**	8.62%
Cerebellum White Matter	**56.65**%	13.90%	29.45%

(b) $11 \times 11 \times 11$ block

	Not Cerebellum	Cerebellum Cortex	Cerebellum White Matter
Not Cerebellum	**77.42%**	15.52%	7.05%
Cerebellum Cortex	**46.67%**	**46.93%**	6.40%
Cerebellum White Matter	**76.49%**	9.94%	13.57%

Sensitivity and specificity were also calculated for this experiment, shown in Table 4. Again, the results show a much lower sensitivity and specificity for the grey level based segmentation than for the previous approach using visual words. Specificity are relatively high (70–90%) for white matter and cortex but in these cases the sensitivity is very low. The opposite is true for the non cerebellum pixels, that are the majority classes with a low specificity and and reasonable sensitivity. This means that grey levels alone are not sufficient for separating the three classes and basically misses out on much of the cortex and white matter classifying this into the majority class of the background.

Table 4. Specificity and sensitivity results for each class with the gray level histogram configuration

(a) Not cerebellum

Block Size	Specificity	Sensitivity
7x7x7	61.24%	61.31%
11x11x11	50.47%	77.42%

(b) Cerebellum cortex

Block Size	Specificity	Sensitivity
7x7x7	75.92%	54.54%
11x11x11	84.54%	46.93%

(c) Cerebellum white matter

Block Size	Specificity	Sensitivity
7x7x7	86.02%	29.45%
11x11x11	93.01%	13.57%

Figure 7 shows the comparison of visual words and gray level histograms in terms of class–wise accuracy for the same block sizes. The results show that for non cerebellum pixels the accuracies differ only slightly between the two techniques, so classification with grey levels alone works reasonably well for this majority class. On the other hand for both cortex and and white matter the performance difference is enormous and these are the two classes that we are most interested in. This underlines that grey level information alone is not sufficient for this classification problem and models or texture information are required for obtaining a good performance.

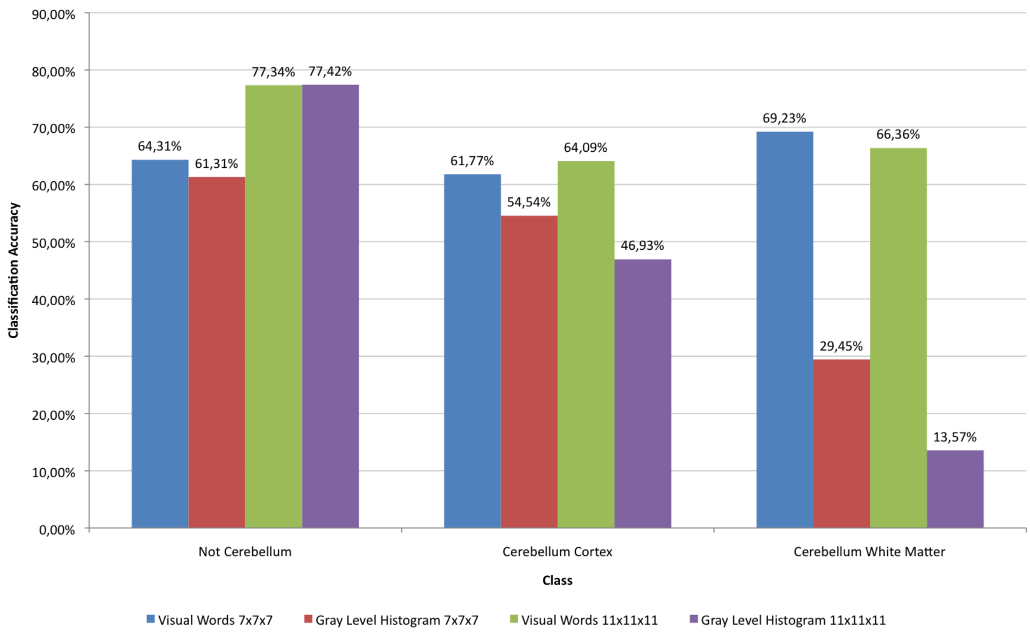


Figure 7. Comparison of classification accuracy using visual words and gray level histograms.

4. DISCUSSION

The results obtained for the segmentation with the various techniques and settings show that our multi–scale visual word approach has a good performance for the task in terms of accuracy, specificity and sensitivity. As can be seen in Figure 6 and Tables 1 and 2, there is an important performance gain when the features are computed in larger block sizes. This shows that texture information in a relatively large neighborhood is important for class membership. Figure 6 shows that this improvement is mostly due to the reduction of the false positives for both cerebellum classes within the not–cerebellum region. These are actually the classes most interesting for this classification and underlines the importance of using a larger neighborhood for the analysis. For the cerebellum cortex and white matter classes, best accuracy values are 64.09% and 69.23%; best specificity values are 83.91% and 92.90% and best sensitivity values are 64.09% and 69.23%.

When compared to the gray level histograms alone, the multi–scale visual words obtain much better results for the cerebellum white matter, which is often confused with the not cerebellum class with using only grey level information. For the cerebellum cortex and not cerebellum classes, results are somewhat similar or slightly better to the ones obtained by the gray level histograms. Still, also for the cerebellum cortex there are fewer errors when using texture information, although there are also more false positives. In our case the cost for errors would be higher than for false positives, so also here the texture information leads to better results. This indicates that the texture patterns in the cerebellum white matter are not rich enough to be distinguished by gray level histograms alone. The multi–scale visual words profit from the fact that the vocabulary is defined based on the patterns that actually occur in the data set and thus are able to distinguish subtle textures patterns. No comparison with a model–based segmentation was attempted in this case but only using simple texture features for the segmentation. Model–based segmentation might lead to slightly better performance but requires the manual creation of a model whereas our approach creates the patterns automatically from the training data. More training data could also potentially increase the performance of our approach. Such an approach could be important in case where subjects differ strongly from a model or when little information is available for model creation as in the case of images from newborns.[26]

5. CONCLUSION

In this article we investigate the use of texture classification using visual words to complement existing techniques for the segmentation of the cerebellum in MR images of the brain. Brain textures are characterized in terms of multi–scale visual words using the wavelet transform as basic visual feature.

Results in Section 3 show that the multi–scale visual words can learn texture patterns of various structures in the brain from the brain MRI images of the cerebellar region. This can complement classical brain structure segmentation methods that use brain models for difficult cases where the shape strongly differs from normal template or atlas (e.g. newborn or pathological brains). Block sizes need to be relatively large to allow for a good characterization of local information to classify the different tissue types. This means that a larger neighborhood of the pixels is required to determine their type.

This article includes several contributions. The possibility of using texture information for segmentation where atlas–based algorithms may fail was investigated. By testing the classifier without any prior spatial knowledge, it was proven that texture features can be valuable for complementing existing segmentation techniques. The introduction of multi–scale visual words by combining the multi–scale coefficients from the wavelet transforms with the bag–of–visual–words approach shows to be adapted to the given problem of cerebellum classification. This contribution can be used as a first step to identify structure in other clinical applications such as similarity–based image search for clinical decision support (i.e., using content–based image retrieval (CBIR)). Once the broad structures are identified, local similarity can help to find similar cases.

The work described also has still several shortcomings that we are currently working on. The optimal block size still needs to be determined and possible the fusion of several scales could lead to optimal results, as border regions of the three tissue classes are expected to be most difficult to characterize, particularly with a growing block size. Thus for border regions small blocks might be better compared to central regions of the three structures. Information of local neighborhoods of pixels could also help in the classification and potentially a simple closing operation could increase the performance. More sophisticated methods for local voting approaches could lead to even better results as could an increased size of the training data. For standard brain segmentation our approach might not be able to increase the performance compared to approaches using cerebellum models but in specific domains such as newborn brain images, such an improvement seems possible as template–based methods do not work well. We are thus currently working on using our techniques on a database of newborn images.

6. ACKNOWLEDGMENTS

This work was partially supported by the Swiss National Science Foundation (FNS) in the MANY project (grant 205321–130046), the EU 7^{th} Framework Program in the context of the Khresmoi project (FP7–257528), and the Center for Biomedical Imaging (CIBM).

REFERENCES

[1] Reinhold Haux. Hospital information systems — past, present, future. *International Journal of Medical Informatics*, 75:268–281, 2005.

[2] Usha Sinha, Alex Bui, Ricky Taira, John Dionisio, Craig Morioka, David Johnson, and Hooshang Kangarloo. A review of medical imaging informatics. *Annals of the New York Acedemy of Sciences*, 980:168–197, December 2002.

[3] Henning Müller, Nicolas Michoux, David Bandon, and Antoine Geissbuhler. A review of content–based image retrieval systems in medicine–clinical benefits and future directions. *International Journal of Medical Informatics*, 73(1):1–23, 2004.

[4] Constantino Carlos Reyes-Aldasoro and Abhir Bhalerao. Volumetric texture segmentation by discriminant feature selection and multiresolution classification. *IEEE Transactions on Medical Imaging*, 26(1):1–14, January 2007.

[5] Naoki Kodama, Yasuhiro Kawase, and Koichiro Okamoto. Application of texture analysis to differentiation of dementia with lewy bodies from alzheimer's disease on magnetic resonance images. In *World Congress on Medical Physics and Biomedical Engineering 2006*, pages 1444–1446. Springer, August 2007.

[6] T. L. Jernigan, S. L. Archibald, C. Fennema-Notestine, A. C. Gamst, J.C. Stout, J. Bonner, and J. R. Hesselink. Effects of age on tissues and regions of the cerebrum and cerebellum. *Neurobiology of Aging*, 22(4):581–594, 2001.

[7] E. W. Y. Tam, D. M. Ferriero, D. Xu, J. I. Berman, D. B. Vigneron, and S. P. Barkovich, A. J. andMiller. Cerebellar development in the preterm neonate: Effect of supratentorial brain injury. *Pediatric Research*, 66(1):102–106, 2009.

[8] B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. Van Der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, and A. M. Dale. Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3):341–355, 2002.

[9] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based tissue classification of mr images of the brain. *IEEE Transactions on Medical Imaging*, 18(10):897–908, 1999.

[10] F. van der Lijn, M. de Bruijne, Y.Y. Hoogendam, S. Klein, R. Hameeteman, M.M.B. Breteler, and W.J. Niessen. Cerebellum segmentation in mri using atlas registration and local multi-scale image descriptors. In *Biomedical Imaging: From Nano to Macro, 2009. ISBI '09. IEEE International Symposium on*, pages 221–224, July 2009.

[11] N. Saeed and B. K. Puri. Cerebellum segmentation employing texture properties and knowledge based image processing: applied to normal adult controls and patients. *Magnetic Resonance Imaging*, 20(5):425–429, 2002.

[12] Hemant D. Tagare, C. Jaffe, and James Duncan. Medical image databases: A content–based retrieval approach. *Journal of the American Medical Informatics Association*, 4(3):184–198, 1997.

[13] Adrien Depeursinge, Alejandro Vargas, Alexandra Platon, Antoine Geissbuhler, Pierre-Alexandre Poletti, and Henning Müller. 3D case–based retrieval for interstitial lung diseases. In *MCBR–CDS 2009: Medical Content–based Retrieval for Clinical Decision Support*, Lecture Notes in Computer Science (LNCS), pages 39–48. Springer, February 2010.

[14] M. T. Suzuki, Y. Yoshitomo, N. Osawa, and Y. Y. Sugimoto. Classification of solid textures using 3D mask patterns. In *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, volume 7, pages 6342–6347, October 2004.

[15] Xin Zhou, Adrien Depeursinge, Richard Stern, Christian Lovis, and Henning Müller. Case–based visual retrieval of fractures. In *International Journal of Computer Assisted Radiology and Surgery*, volume 5, Supplement 1, pages 11548/S162–S163, Geneva, Switzerland, June 2010. Springer.

[16] Adrien Depeursinge, Alejandro Vargas, Frédéric Gaillard, Alexandra Platon, Antoine Geissbuhler, Pierre-Alexandre Poletti, and Henning Müller. Case–based lung image categorization and retrieval for interstitial lung diseases: clinical workflows. *International Journal of Computer Assisted Radiology and Surgery*, pages 1–14, 2011.

[17] G. Quellec, M. Lamard, G. Cazuguel, C. Roux, and B. Cochener. Case retrieval in medical databases by fusing heterogeneous information. *IEEE Transactions on Medical Imaging*, 30(1):108–118, jan. 2011.

[18] Adrien Depeursinge and Henning Müller. Fusion techniques for combining textual and visual information retrieval. In Henning Müller, Paul Clough, Thomas Deselaers, and Barbara Caputo, editors, *ImageCLEF*, volume 32 of *The Springer International Series On Information Retrieval*, pages 95–114. Springer Berlin Heidelberg, 2010.

[19] Antonio Foncubierta-Rodríguez, Adrien Depeursinge, and Henning Müller. Using multiscale visual words for lung texture classification and retrieval. In Hayit Greenspan, Henning Müller, and Tanveer Syeda Mahmood, editors, *Medical Content-based Retrieval for Clinical Decision Support*, volume 7075 of *MCBR-CDS 2011*. Lecture Notes in Computer Sciences (LNCS), September 2011.

[20] Uri Avni, Hayit Greenspan, Eli Konen, Michal Sharon, and Jacob Goldberger. X–ray categorization and retrieval on the organ and pathology level, using patch–based visual words. *IEEE Transactions on Medical Imaging*, 30(3):733–746, 2011.

[21] Michael Unser. Texture classification and segmentation using wavelet frames. *IEEE Transactions on Image Processing*, 4(11):1549–1560, November 1995.

[22] Josef Sivic and Andrew Zisserman. Video google: A text retrieval approach to object matching in videos. In *Proceedings of the Ninth IEEE International Conference on Computer Vision - Volume 2*, ICCV '03, pages 1470–1477, Washington, DC, USA, 2003. IEEE Computer Society.

[23] C. J. van Rijsbergen. *Information Retrieval*. Prentice Hall, Englewood Cliffs, New Jersey, USA, 1979.

[24] J. Iivarinen and A. Visa. An adaptive texture and shape based defect classification. In *Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on*, volume 1, pages 117–122, aug 1998.

[25] A.P. Dhawan, Y. Chitre, and C. Kaiser-Bonasso. Analysis of mammographic microcalcifications using gray-level image structure features. *Medical Imaging, IEEE Transactions on*, 15(3):246–259, jun 1996.

[26] Laura Gui, Radoslaw Lisowski, Tamara Faundez, Petra S. Huppi, François Lazeyras, and Michel Kocher. Automatic segmentation of newborn brain MRI using mathematical morphology. In *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 2026–2030, april 2011.