

Sulci Detection in Photos of the Human Cortex Based on Learned Discriminative Dictionaries

Benjamin Berkels¹, Marc Kotowski², Martin Rumpf³, and Carlo Schaller²

¹ Interdisciplinary Mathematics Institute,
University of South Carolina, Columbia, SC 29208, USA,
`berkels@mailbox.sc.edu`,

² Hôpitaux Universitaires de Genève,
Rue Micheli-du-Crest 24, 1211 Genève, Switzerland

³ Institut für Numerische Simulation,
Rheinische Friedrich-Wilhelms-Universität Bonn,
Endenicher Allee 60, 53115 Bonn, Germany
`martin.rumpf@ins.uni-bonn.de`

Abstract The use of discriminative dictionaries is exploited for the segmentation of sulci in digital photos of the human cortex. Manual segmentation of the geometry of sulci by an experienced physician on training data is taken into account to build pairs of such dictionaries. It is demonstrated that this approach allows a robust segmentation of these brain structures on photos of the brain as long as the training data contains sufficiently similar images. Concerning the methodology an improved minimization algorithm for the underlying variational approach is presented taking into account recent advances in orthogonal matching pursuit. Furthermore, the method is stable since it ensures an energy decay in the dictionary update.

1 Introduction

In neurosurgery, a major challenge is the adaption of pre-surgery acquired brain images and cortex geometry to the intra-interventional brain configuration. Digital photos can be easily taken through the microscope and provide information on the currently observed brain shift. Sulci are the most prominent geometric characteristics visible on such photos. As illustrated by Figure 1, the detection of sulci in such images is a very challenging task. For instance, some of the sulci are covered by blood vessels while the very same blood vessels also cover part of the gyri. Therefore, pixelwise segmentation approaches based on the color values cannot be sufficient to handle this segmentation problem, not even when color distributions learned from images manually marked by an expert are used. In this paper, we use the concept of learned discriminative dictionaries to segment the geometry of sulci in 2D digital photos. Thereby, on a training data set an experienced physician marks the sulci geometry, which will then be used to build a suitable discriminative dictionary.

Nowadays, sparse signal representations based on overcomplete dictionaries are used for a wide range of signal and image processing tasks. The key assumption of these models is that finite dimensional signals can be well approximated by sparse linear combinations of so-called *atoms* or *atom signals*. Due to their finite dimensionality, the signals and the atoms are considered to be elements of \mathbb{R}^N . A set of atoms d_1, \dots, d_K is called *dictionary* and represented by the matrix $D \in \mathbb{R}^{N \times K}$ whose j -th column is the atom d_j .

There are two main variants of the sparse approximation problem, the *error-constrained* approaches and the *sparsity-constrained* approaches. Here, we are considering an approach of the latter type: For a given input signal y (in our application a patch from a digital photo of the brain) we ask for its best approximation under the constraint that at most $L \in \mathbb{N}$ atoms are used, i. e.

$$\min_{x \in \mathbb{R}^K} \|y - Dx\|^2 \text{ such that } \|x\|_0 \leq L,$$

where $\|\cdot\|_0$ denotes the l_0 “norm”, i. e. the number of nonzero components.

One of the major challenges in the context of sparse representations is the design of suitable dictionaries. The sparse representation itself usually is just a means to an end and used to solve a certain task like, for instance, denoising or compression. Thus, the dictionary has to be tailored to the actual imaging task. In general, there are two distinct approaches to dictionary design: The simpler and more traditional route is to use a predefined dictionary generated by a transform like the short-time Fourier transform [2], the wavelet [12], curvelet [6] or contourlet transform [8], to name just a few. The more sophisticated approach is to learn the dictionary from the input data or some representative training data. A very popular and highly efficient representative of this kind is the K-SVD algorithm [1].

Like K-SVD, most of the existing dictionary learning algorithms aim at generating *reconstructive* dictionaries, i. e. dictionaries that are optimized to sparsely represent a certain class of input signals or images. In this paper, our goal is to detect sulci on the human cortex in digital photographs. Thus, we need to distinguish between different types of signals which gives rise to so-called *discriminative* dictionaries. These kind of dictionaries not only aim to give a suitable representation of a given type of signals, but are also optimized to be not as suitable for the reconstruction of a different given class of signals. Mairal et al. [10] introduced a variational approach to learning discriminative dictionaries for local image analysis and presented a multiscale extension applied to class-specific edge detection [11].

Zhao et al. [15] combine the discriminative dictionary model from [10] with additional pre- and post-processing stages to optimize the discriminative approach for text detection in images. Zhang and Li [14] propose a different route to discriminative dictionary learning: They extend the K-SVD algorithm to solve for a dictionary and a classifier simultaneously and claim that this kind of algorithm is less likely to get stuck in local minima than the one from [10]. Let us remark here, that their K-SVD extension still uses an alternating minimiza-

tion scheme to solve a non-convex minimization problem. Hence, there is no guarantee that the global optimum is finally found.

The contributions of this paper are twofold: On the one hand, we introduce an improved minimization algorithm for the variational approach to discriminative dictionaries from [10]. This algorithm is more efficient because it incorporates recent advances in orthogonal matching pursuit made by Rubinstein et al. [13] and it is more stable since it ensures an energy decay in the dictionary update unlike the truncated Newton iteration used in [10,11]. On the other hand, we study the applicability of discriminative dictionaries to detect sulci on the intra-operative digital photographs of the human cortex. As we will see in this paper, manually marked images can indeed be used to learn a discriminative dictionary pair and thereby allow to detect sulci on images as long as the training data contains sufficiently similar brain images.

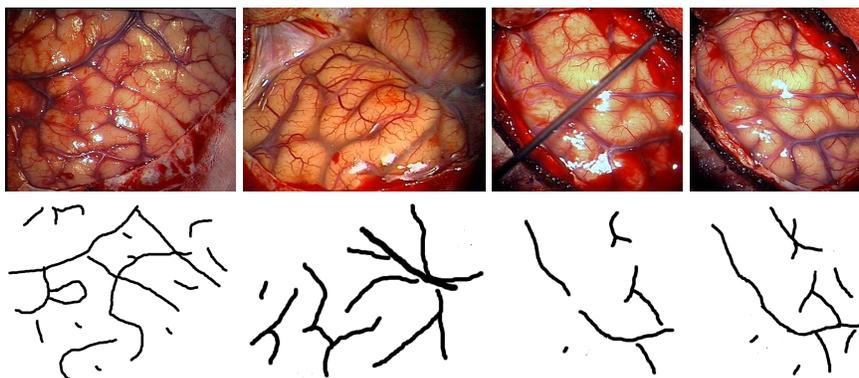


Figure 1. Four typical digital photographs of the exposed human cortex (top row) and the sulci regions of the cortex manually marked by an expert (bottom row).

2 Learning Discriminative Dictionaries

Given M input patches $y_1, \dots, y_M \in \mathbb{R}^N$, a reconstructive dictionary tailored to these patches can be learned with the minimization problem

$$\min_{X \in \mathbb{R}^{K \times M}, D \in \mathbb{R}^{N \times K}} \sum_{l=1}^M R(y_l, D, x_l) \text{ such that } \|x_l\|_0 \leq L \text{ for } 1 \leq l \leq M.$$

Here, x_l denotes the l -th column of X and $R(y, D, x) := \|y - Dx\|^2$ is the reconstruction error of a patch $y \in \mathbb{R}^N$ for a dictionary $D \in \mathbb{R}^{N \times K}$ and dictionary coefficients $x \in \mathbb{R}^K$. Well-known algorithms to tackle this minimization problem are the method of optimal directions (MOD) [9] or K-SVD [1].

Denoting the coefficients of the sparse best approximation of y using D by

$$x^*(y, D) := \operatorname{argmin}_{x \in \mathbb{R}^K, \|x\|_0 \leq L} R(y, D, x),$$

the best approximation error is $\mathcal{R}(y, D) := R(y, D, x^*(y, D))$. Then the minimization problem for reconstructive dictionary learning is equivalent to

$$\min_{D \in \mathbb{R}^{N \times K}} \sum_{l=1}^M \mathcal{R}(y_l, D). \quad (1)$$

With this notation we can formulate the discriminative dictionary approach of Mairal et al. [10]. Since our application, the detection of sulci on the human cortex, only requires two labels, we here explicitly formulate only the two label case. The extension to multiple labels is straightforward and our algorithm can be easily adapted to more than two labels.

Given input patches $y_1, \dots, y_{M_1+M_2}$ of two different classes P_1 and P_2 , where $P_i := \{y_l : l \in S_i\}$, $S_1 = \{1, \dots, M_1\}$ and $S_2 = \{M_1 + 1, \dots, M_1 + M_2\}$, a pair of discriminative dictionaries can be found solving the minimization problem

$$\min_{D_1, D_2} \sum_{i=1}^2 \frac{1}{M_i} \sum_{l \in S_i} [C_\lambda((-1)^{i+1}(\mathcal{R}(y_l, D_2) - \mathcal{R}(y_l, D_1))) + \lambda\gamma\mathcal{R}(y_l, D_i)]. \quad (2)$$

Here, C_λ denotes the logistic loss function, i. e. $C_\lambda(s) = \ln(1 + \exp(-\lambda s))$, and λ, γ are nonnegative constant parameters. The last summand is already known from the reconstructive learning problem (1) and handles the reconstructive properties of our dictionary pair. The first summand is responsible for the discriminative properties of the dictionaries. For instance, for $i = 1$ and $l \in S_1$ we have

$$C_\lambda((-1)^2(\mathcal{R}(y_l, D_2) - \mathcal{R}(y_l, D_1))) \begin{cases} \approx 0 & \mathcal{R}(y_l, D_1) \ll \mathcal{R}(y_l, D_2) \\ \gg 0 & \mathcal{R}(y_l, D_1) \gg \mathcal{R}(y_l, D_2). \end{cases} \quad (3)$$

In other words, this logistic loss term is small, if and only if D_1 is more suitable to reconstruct P_1 (the signals from the first class) than D_2 is.

3 Minimization Algorithm

The discriminative minimization problem (2) is highly nonconvex and requires a carefully chosen numerical minimization strategy. Like [10], our minimization strategy is based on the K-SVD algorithm and consists of a *sparse coding stage* and a *codebook update stage*.

In the sparse coding stage, the sparse approximation coefficients are computed for all patches using the current estimates for both dictionaries, i. e. $x_l^i \approx x^*(y_l, D_i)$ for $i = 1, 2$ and $l = 1, \dots, M_1 + M_2$, cf. Algorithm 1.1. Instead of using OMP for this as suggested in [10,11], we propose to use the Batch-OMP algorithm from [13]. This algorithm is based on the fact that the *same* dictionary

is used to code a *large* set of signals. In particular, it exploits the fact that in the atom selection step of OMP neither the residual r nor the coefficients x need to be known, but only $D^T r$. As shown in [13], Batch-OMP is almost an order of magnitude faster than OMP when used on sufficiently many input signals.

In [11], a different way to speed up the algorithm from [10] is proposed: Also noting that the sparse coding stage is computationally expensive, they propose to update the dictionaries and the coefficients with fixed sparsity pattern in the codebook update stage by alternating till convergence instead of doing so only once to reduce the number of sparse coding steps. This idea is complementary to our proposal to speed up the algorithm and thus can be used in combination with it.

In the codebook update stage, the dictionaries and the coefficients are updated while keeping the obtained sparsity pattern fixed during the sparse coding stage. [10,11] propose to do this update with a truncated Newton method. “Truncated” here refers to the fact that this method neglects the second derivatives of C_λ . In our experiments with manually marked images of the human cortex, this algorithm unfortunately had numerical stability problems and didn’t always produce sufficiently discriminative dictionary pairs. This is most likely because the truncated Newton method does not guarantee an energy decay of the target functional due to the lack of an appropriate step size control. Furthermore, C_λ is not approximately linear at 0, the transition between the nearly linear and the nearly constant part of C_λ which is the important transition region between the two cases outlined in (3). Therefore, neglecting of the second derivatives of C_λ is questionable.

To update a single entry of one of the dictionaries, we use a step size controlled gradient descent on the functional from (2) while freezing the coefficients, i. e. $\mathcal{R}(y_l, D_i)$ is approximated by $R(y_l, D_i, x_l^i)$ and thus we use the functional

$$E[D_1, D_2] = \sum_{i=1}^2 \frac{1}{M_i} \sum_{l \in S_i} [C_\lambda ((-1)^{i+1} (R(y_l, D_2, x_l^2) - R(y_l, D_1, x_l^1))) + \lambda \gamma R(y_l, D_i, x_l^i)]$$

and update d_j^1 by $d_j^1 - \tau \partial_{d_j^1} E[D_1, D_2]$ where τ is determined using the *Armijo rule* [3,5]. Note that the specific choice of the step size control is not important here, but it is important to use a step size control that guarantees an energy decay. Like K-SVD, we assume the dictionary entries to be normalized, i. e. $\|d_j^i\| = 1$, and therefore scale the dictionary entry accordingly after the gradient descent update. Using a straightforward calculation one obtains

$$\partial_{d_j^1} E[D_1, D_2] = 2 \sum_{i=1}^2 \sum_{l \in S_i} w_l^i (x_l^1)_j (D_1 x_l^1 - y_l) ,$$

where

$$w_l^i = \frac{1}{M_i} ((-1)^i C'_\lambda ((-1)^{i+1} (R(y_l, D_2, x_l^2) - R(y_l, D_1, x_l^1))) + \delta_{i1} \lambda \gamma) .$$

Denoting the j -th entry of x_l^1 by $(x_l^1)_j$ and using

$$E_l^1[D, j] = (y_l - Dx_l^1 + (x_l^1)_j d_j)$$

as well as the indices of patches that use d_j^1 , i.e. $\omega_j^1 := \{l : (x_l^1)_j \neq 0\}$, the variation can be expressed as

$$\partial_{d_j^1} E[D_1, D_2] = 2 \sum_{i=1}^2 \sum_{l \in S_i \cap \omega_j^1} w_l^i (x_l^1)_j [(x_l^1)_j d_j^1 - E_l^1[D_1, j]].$$

Replacing the sum $\sum_{l \in S_i}$ by $\sum_{l \in S_i \cap \omega_j^1}$ is crucial to keep the computational cost for the codebook update stage within reasonable limits. The same replacement can be done in E when it needs to be evaluated for the Armijo rule. After updating a dictionary entry, we update the corresponding coefficients keeping the sparsity pattern. Similarly to the representation of $\partial_{d_j^1} E$, one now obtains

$$\partial_{(x_l^1)_j} E = \sum_{i=1}^2 w_l^i \partial_{(x_l^1)_j} R(y_l, D_1, x_l^1) = \partial_{(x_l^1)_j} R(y_l, D_1, x_l^1) \sum_{i=1}^2 w_l^i.$$

Therefore, $\partial_{(x_l^1)_j} E = 0$ holds when $\partial_{(x_l^1)_j} R(y_l, D_1, x_l^1) = 0$. Using

$$\partial_{(x_l^1)_j} R(y_l, D_1, x_l^1) = 2 \left((x_l^1)_j \|d_j^1\|^2 - E_l^1[D_1, j] \cdot d_j^1 \right)$$

and $\|d_j^1\|^2 = 1$ leads to the update formula $(x_l^1)_j \leftarrow E_l^1[D_1, j] \cdot d_j^1$. The dictionary D_2 and its corresponding coefficients can be updated analogously.

Like [10], we use an ascending series for the parameter λ and a descending series for γ . Since our codebook update stage is guaranteed not to increase E , we do not need the sophisticated strategy to adaptively adjust the parameters used in [10]. Instead, in all our experiments, we simply used $\lambda = 100k$ and $\gamma = 1/k$ in the k -iteration of the algorithm. A sketch of this computational procedure is given in Algorithm 1.1.

4 Segmentation with Discriminative Dictionaries

For the detection of sulci in images of the human cortex, we assume to be given a number of human cortex images where the sulci were manually marked by a physician. These images are then separated into small patches of a user selectable patch size and the patches are divided into two sets, sulci and non-sulci patches depending on whether the central pixel of the patch belongs to the sulci region marked by the physician. Using these two sets of patches, a discriminative dictionary pair is learned using the Algorithm 1.1.

Using this dictionary pair, images can be segmented into sulci and non-sulci regions using a binary Mumford-Shah model where the reconstruction errors with the two dictionaries are used as the two indicator functions. A global minimizer of this model is calculated using the convex reformulation of the problem proposed in [4] and using [7, Algorithm 2] to efficiently calculate a minimizer of the convex functional.

Algorithm 1.1: General minimization strategy

```

given input patches  $y_1, \dots, y_{M_1+M_2}$  of two different classes  $P_1$  and  $P_2$ ;
initialize  $D_1$  and  $D_2$  with K-SVD from  $P_1$  and  $P_2$  respectively;
initialize  $k = 0$ ;
repeat
   $k \leftarrow k + 1$ ;
   $\lambda = 100k, \gamma = 1/k$ ;
  Sparse coding stage;
  for  $i = 1$  to 2 do
    for  $l = 1$  to  $M_1 + M_2$  do
      Calculate  $x_l^i \approx x^*(y_l, D_i)$  using Batch-OMP;
    end
  end
  Codebook update stage;
  for  $i = 1$  to 2 do
    for  $j = 1$  to  $K$  do
      Calculate  $\omega_j^i = \{l : (x_l^i)_j \neq 0\}$ ;
       $d_j^i \leftarrow d_j^i - \tau \partial_{d_j^i} E$  determining  $\tau$  using the Armijo rule;
       $d_j^i \leftarrow d_j^i / \|d_j^i\|$ ;
      for  $l \in \omega_j^i$  do
         $(x_l^i)_j \leftarrow E_l^i[D_i, j] \cdot d_j^i$ ;
      end
    end
  end
until convergence;

```

5 Results

As first experiment we verify the general applicability of our discriminative dictionary approach to detect sulci on intra-operative images of the human cortex. For this we learn a discriminative dictionary pair from a single manually marked image as described in the previous section. The first row of Figure 2 shows that in this optimal case, the segmentation almost perfectly matches the manual markings made by the expert. Here, we used a patch size of 13×13 , and $K = 256$, $L = 4$ and $\gamma = 0.00002$ as parameters, where γ denotes the weighting of the regularity term in the Mumford–Shah model segmentation model.

In order to cut down the computational time given the fact that there are considerably more non-sulci than sulci input patches, we randomly selected 30% of the non-sulci patches instead of using all of them for the dictionary learning in a second experiment, cf. second row of Figure 2. Indeed, this only slightly reduces the accuracy of the segmentation. Henceforth, we only use 30% of the non-sulci patches to learn the dictionaries in the remaining experiments.

In the next experiment, we use three frames from an intra-operative video, all with sulci manually marked by a physician, to learn a discriminative dictionary pair. We use the same values for K , L and γ as in the previous experiments and

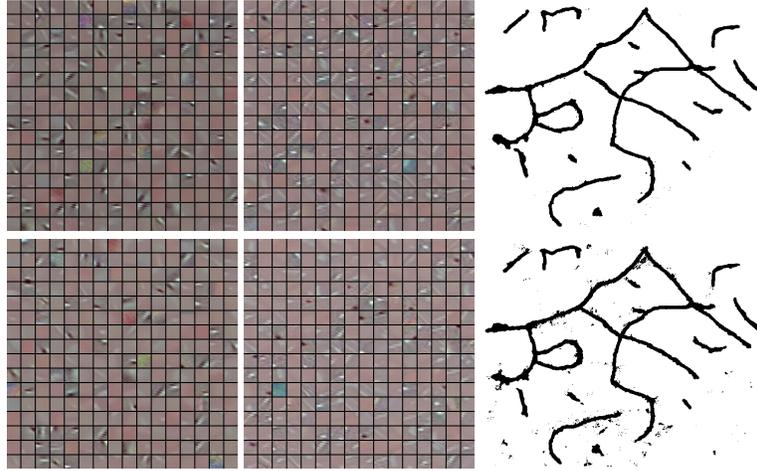


Figure 2. Discriminative dictionary pairs learned from the manually marked image shown on the left of Figure 1 and segmentation of this image based on these dictionary pairs. In the top row all available non-sulci patches were used to learn the dictionaries, while in the bottom row only 30% of the non-sulci patches were used.

a patch size of 12×12 and 20×20 . Figure 3 shows the resulting dictionaries while Figure 4 shows the segmentation obtained with these dictionaries. Segmentation on the frames already used in the learning phase is almost perfect. There are only a few minor artifacts compared to the manual segmentation performed by the physician. Although not surprising, this confirms that a discriminative dictionary pair has no problems encoding information from multiple input frames. It can also be seen from this figure that increasing the patch size from 12×12 to 20×20 slightly improves the results. The second row of Figure 4 is more interesting: It shows that the dictionaries can also be used to segment frames that were not used during the learning process. The dictionary based segmentation shows some artifacts away from the cortex region, but this is due to the fact that images used to learn the dictionaries were cropped to the cortex region and thus the dictionaries cannot contain information about these areas. This kind of effect can also be seen inside the cortex region: On the top right of the manual markings of the physician for this image is a small sulci that is not found in the dictionary based segmentation. This is just natural since the physician did not mark this sulci in the frames that were used for the learning, cf. top row of Figure 4.

In the final set of experiments, we use thirteen manually marked images, the three images already used in the previous experiment and ten from another intra-operative video to learn discriminative dictionaries. We use the same values for K , L and γ as in the previous experiments and a patch size of 13×13 and 19×19 . Figure 5 shows the resulting dictionaries while Figure 6 shows the segmentation obtained with these dictionaries on multiple frames. The first three images were

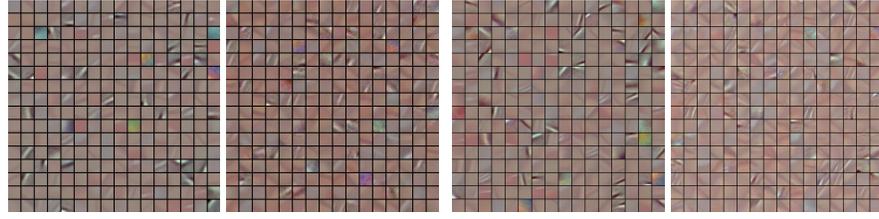


Figure 3. Discriminative dictionary pairs with patch size of 12×12 (left pair) and 20×20 (right pair) learned from three frames of an intra-operative video.



Figure 4. Two cortex images (first column), manual segmentation of the sulci by an expert (second column) and segmentation obtained using the dictionaries from Figure 3 of patch size 12×12 (third column) and 20×20 (forth column). Note that the manual marking from the top row was used during the dictionary learning but the one from the bottom row was not.

used in the dictionary learning process, so it comes as no surprise that the obtained segmentations are close to the manual markings. One observation here is remarkable though. In the image shown in the third row, the physician did not mark the sulci in the lower right part of the image even though he marked that sulci in other frames of the same video sequence, for instance in the frame shown in the second row. Nevertheless, the dictionary based segmentation is able to detect traces of these sulci because underlying information is encoded of more than just the marking of this single image and thus the method can average out conflicting markings.

The remaining four rows show results of the dictionary based segmentation on images that were not used while learning the dictionaries. While the segmentation understandably is not as good on these frames as on the frames used during the learning, the sulci structures are still clearly identified in the forth to sixth row. Here, it is also evident that increasing the patch size from 13×13 to 19×19 has a positive effect on the quality of the segmentation. With the

larger patch size the width of the detected sulci is more accurate and there are less artifacts in the non-sulci regions. Finally, in the last row, the limits of the dictionary based segmentation approach become visible. The cortex region in this frame differs too much from the cortex regions in the learning frames and thus the sulci are not properly detected here.

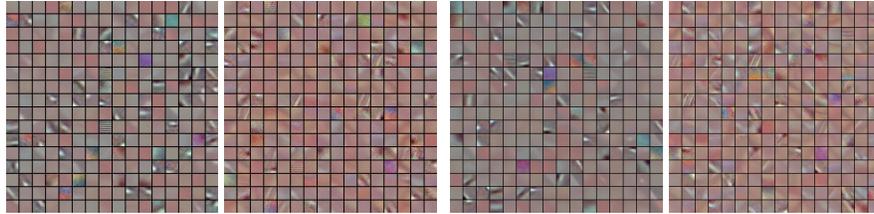


Figure 5. Discriminative dictionary pairs with patch size of 13×13 (left pair) and 19×19 (right pair) learned from a total of thirteen different frames originating from two different intra-operative videos.

6 Conclusion

We have studied the applicability of discriminative dictionaries to segment the geometry of sulci in intra-operative digital photographs of the human cortex. It turned out that human cortex images manually marked by an experienced physician can be used to learn discriminative dictionary pairs that allow for a robust segmentation of these brain structures on photos of the cortex as long as the training data contains sufficiently similar images.

Furthermore, we have presented an improved minimization strategy for the discriminative dictionary functional of Mairal et al. [10] that is more efficient by leveraging recent advances in orthogonal matching pursuit and more stable due to a new dictionary update step that ensures an energy decay.

References

1. Aharon, M., Elad, M., Bruckstein, A.: K-SVD: An algorithm for designing over-complete dictionaries for sparse representation. *IEEE Transactions on Signal Processing* 54(11), 4311–4322 (November 2006)
2. Allen, J.B.: Short term spectral analysis, synthesis, and modification by discrete fourier transform. *IEEE Transactions on Acoustics, Speech and Signal Processing ASSP-25(3)*, 235–238 (June 1977)
3. Armijo, L.: Minimization of functions having Lipschitz continuous first partial derivatives. *Pacific Journal of Mathematics* 16(1), 1–3 (1966)

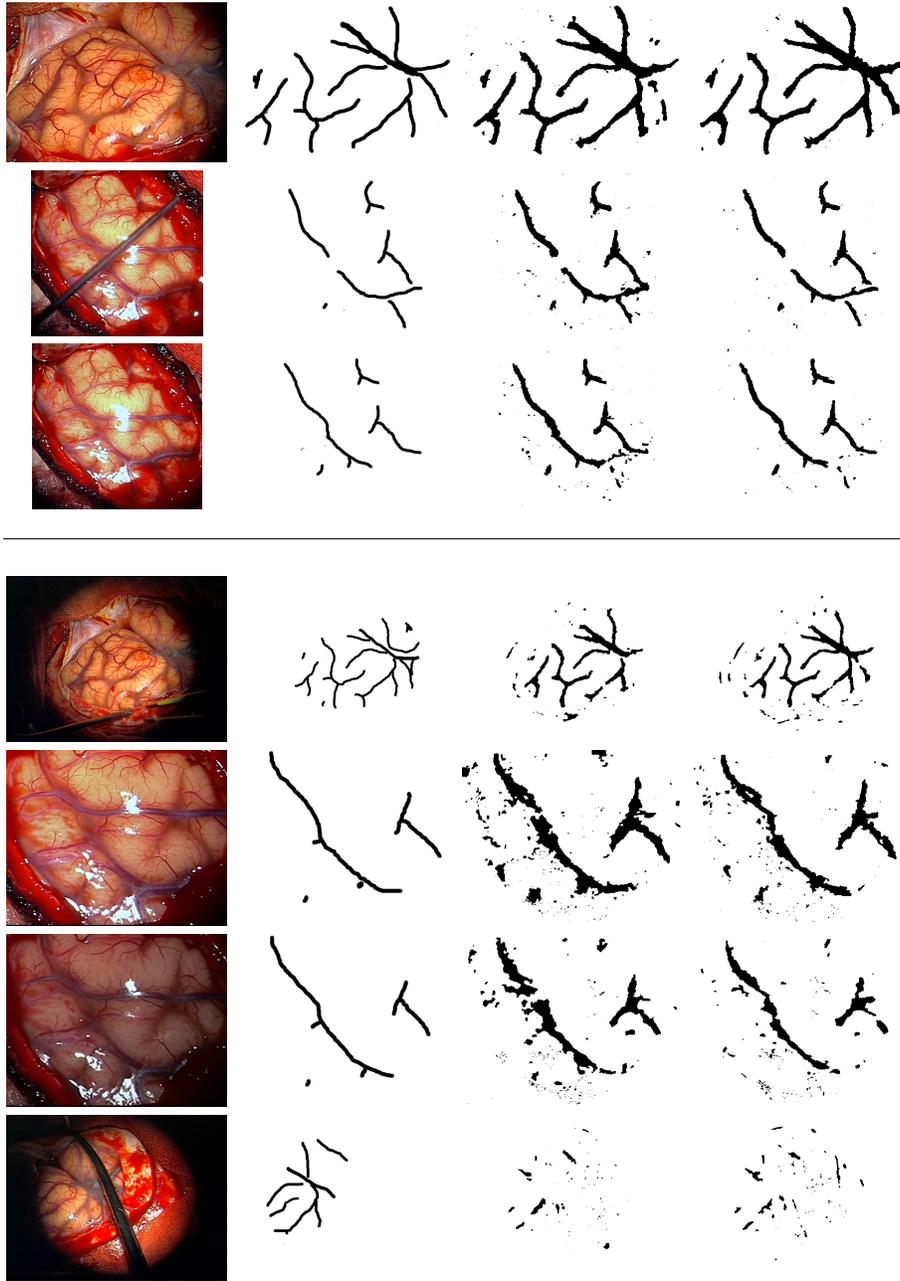


Figure 6. Multiple cortex images (first column), manual segmentation of the sulci by an expert (second column) and segmentation obtained using the dictionaries from Figure 5 of patch size 13×13 (third column) and 19×19 (forth column). Note that the manual markings from the first three rows were used during the dictionary learning but the markings from the other rows were not.

4. Berkels, B.: An unconstrained multiphase thresholding approach for image segmentation. In: Proceedings of the Second International Conference on Scale Space Methods and Variational Methods in Computer Vision (SSVM 2009). Lecture Notes in Computer Science, vol. 5567, pp. 26–37. Springer (2009)
5. Bertsekas, D.P.: Nonlinear Programming. Athena Scientific, Belmont, MA, 2nd edn. (1999)
6. Candès, E.J., Donoho, D.L.: Curvelets – a surprisingly effective nonadaptive representation for objects with edges. In: Schumaker, L.L., et al. (eds.) Curves and Surfaces. Vanderbilt University Press, Nashville, TN (1999)
7. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. Tech. Rep. 685, Ecole Polytechnique, Centre de Mathématiques appliquées, UMR CNRS 7641, 91128 Palaiseau Cedex (France) (May 2010)
8. Do, M.N., Vetterli, M.: The contourlet transform: an efficient directional multiresolution image representation. IEEE Transactions on Image Processing 14(12), 2091–2106 (December 2005)
9. Engan, K., Aase, S.O., Husøy, J.H.: Frame based signal compression using method of optimal directions (MOD). In: Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS'99). vol. 4, pp. 1–4 (1999)
10. Mairal, J., Bach, F., Ponce, J., Sapiro, G., Zisserman, A.: Discriminative learned dictionaries for local image analysis. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1–8 (2008)
11. Mairal, J., Leordeanu, M., Bach, F., Hebert, M., Ponce, J.: Discriminative sparse image models for class-specific edge detection and image interpretation. In: Computer Vision – ECCV 2008. LNCS, vol. 5304, pp. 43–56 (2008)
12. Mallat, S.: A wavelet tour of signal processing. Academic Press (1999)
13. Rubinstein, R., Zibulevsky, M., Elad, M.: Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit. Tech. rep., CS Technion (April 2008)
14. Zhang, Q., Li, B.: Discriminative K-SVD for dictionary learning in face recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2691–2698 (2010)
15. Zhao, M., Li, S., Kwok, J.: Text detection in images using sparse representation with discriminative dictionaries. Image and Vision Computing 28(12), 1590–1599 (December 2010)