The final publication is available at

https://doi.org/10.1016/j.image.2017.03.018

Additional Information

Assessment of Sparse-based Inpainting for Retinal Vessel Removal

A. Colomer[a,b,*], V. Naranjo[a,b], K. Engan[c], K. Skretting[c]

[a]*Instituto de Investigación e Innovación en Bioingeniería (I3B), Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia, Spain.*
[b]*Grupo Tecnologías de Informática Aplicadas a la Oftalmología, Unidad Conjunta UPV-FISABIO, Spain.*
[c]*Department of Electrical Engineering and Computer Science, University of Stavanger, N-4036 Stavanger, Norway.*

Abstract

Some important eye diseases, like macular degeneration or diabetic retinopathy can induce changes visible on the retina, for example as lesions. Segmentation of lesions, or extraction of textural features from the fundus images are possible steps towards automatic detection of such diseases which could facilitate screening as well as provide support for clinicians. For the task of detecting significant features, retinal blood vessels are considered as being interference on the retinal images. If these blood vessel structures could be suppressed, it might lead to a more accurate segmentation of retinal lesions as well as a better extraction of textural features to be used for pathology detection. This work proposes the use of sparse representations and dictionary learning techniques for retinal vessel inpainting. The performance of the algorithm is tested for grayscale and RGB images from the DRIVE and STARE public databases, employing different neighbourhoods and sparseness factors. Moreover, a comparison with the most common inpainting family, diffusion-based methods, is carried out. For this pourpose, two different ways of assessing the quality of the inpainting are presented and used to evaluate the results of the non-artifical inpainting, i.e. where a reference image does not exist. The results suggest that the use

---

*Corresponding author
   Email address: `adcogra@i3b.upv.es` (A. Colomer)

of sparse-based inpainting performs very well for retinal blood vessels removal which will be useful for the future detection and classification of eye diseases.

1. Introduction

Age-related macular degeneration (AMD) and diabetic retinopathy (DR) are nowadays two of the most frequent causes of blindness and vision impairment in the world [1]. The early diagnosis of these pathologies is very important. However, due to the large of population at risk, a potential screening would be highly beneficial for clinicians.

Two examples of fundus images with lesions are shown in Figure 1. Blood vessels cover a high percentage of the fundus image hindering the automatic detection of important structures as optic disc, optic cup and macula among others. Vessels are also considered as noise or artefacts that hamper the segmentation of different lesions such as exudates, microaneurysms and drusen among others (Figure 1) and also the classification of pathologies based on background textures. A possible procedure to avoid blood vessels is to consider these structures as missing pixels and trying to restore them using the background.

Figure 1: (a) Exudates, microaneurysms and (b) drusen in pathological fundus images.

2

Image inpainting is a technique for restoring missing or damaged areas in an digital image. Image inpainting has been recently used in a lot of applications such as in image restoration from scratches or text overlays, for removal or replacement of selected objects, for disocclusion in image-based tracking in order to recover missing persons or objects in specific frames, etc.

Inpainting methods assume that pixels in the known and unknown areas of the image have the same statistical properties or geometrical structures. Different algorithms exist in the literature but all of them can be grouped in three categories or families.

Diffusion-based inpainting covers a family of methods exploiting the smoothness principle by parametric models or partial differential equations (PDEs) to propagate local structures from the exterior to the interior of the unknown area. These algorithms are designed for completing straight lines and curves, for processing piecewise smooth images, for propagating strong structures and filling small gaps. Using this kind of methods to fill large gaps produce a blur result [2]. Different state-of-the-art approaches exist using particular models (linear, non-linear, isotropic or anisotropic) to favour the propagation in specific directions [3–5].

The second category is examplar-based inpainting that exploits image statistical and self-similarity principles. Using methods belonging at this category, the unknown region is filled by means of copying and stitching the best match pixel or patch from the sample texture. Exemplar-based methods follow two different ways. In pixel-based algorithms the unknown area can be filled copying the central pixel of the most similar patch from the sample texture. The idea of patch-based algorithms is to fill the unknown region in one step by sampling and copying texture patterns (entire patches) from the source [6–8]. Obviously, patch-based solutions are more computationally efficient than pixel-based approaches, for this reason nowadays exemplar-based methods refers to patch-based methods.

Different researchers focus their efforts on improving and optimizing patch-based methods trying to find the most efficient way to search the best match

3

patch. One of them are the sparse-based algorithms [9, 10] where the image is supposed to be sparse in a given domain and known and unknown parts are assumed to share the same sparse representation. The domain can be represented by a dictionary matrix which can be learned using dictionary-learning methods.

The last category, hybrid methods, are based on separating the image information (texture and structure), and thereafter inpaint these components separately using the most appropriate method (diffusion-based inpainting or examplar-based inpainting).

Another important line of research closely related to image inpainting is the quality assesment of the resulting image [11–13]. The efforts are focused on looking for an objective measure that predict automatically the human perception of an inpainted image without the human help. When performing inpainting on real world data without a reference image available (non-articial inpainting), a quality assessment of the inpainting is not straight forward. In this paper two different ways to measure the quality for non-artificial inpainting are proposed and validated in a real application.

Different authors [14–16] have inpainted blood vessels by using diffusion-based algorithms but only as a middle step of their purpose, in other words, without any kind of evaluation of the inpainting method used. For example, [15] uses inpainting techniques in order to segment the optic disk in fundus images. Inpainting quality evaluation, in this application, is an unsolved problem because the reference image (fundus image without vessels) does not exist. Image inpainting assessment is an essential step for our future purposes because the better the vessels are removed, the more accurate the texture classification or the lesion detection will be.

The main novelty of this work is to inpaint blood vessels in fundus images using a sparse-based inpainting method with spread neighbourhoods and to establish a way to asses inpainting methods.

In this work, three experiments were carried out in order to find the best configuration of the proposed inpainting method applied to retinal vessel removal. The influence of using spread neighbourhoods, the variation of the sparseness

4

factor, the algorithm used in the dictionary learning process and the way to build the RGB dictionaries are some of the key concepts studied in these experiments.

In this work, the blood vessels of grayscale and color retinal images, from two public databases, were inpainted. The resulting images were evaluated using two metrics. Firstly, the traditional Recovery Error is found on some artificially made vessels added to the image. In this paper we also present a novel inpainting quality assessment index based on the idea exposed in [17]. By means of this index, a comparison between our sparse-based method and a diffusion-based inpainting algorithm, via partial differential equations (PDE), is carried out.

The rest of the paper is organized as follows: in Section 2, the inpainting methods are described and the metrics for measuring the quality of non-artificial inpainted images are introduced. Section 3 shows the experimental results of blood vessels inpainting in fundus images and the assessment of each inpainting algorithm and configuration using the proposed metrics. Finally, section 4 provides conclusions and some areas for future work.

2. Methods

2.1. Sparse-Based Inpainting

A missing pixel region can be filled using methods that synthesize entire patches by learning from patches in the known part of the image [2]. Different variants in the fast searching of best matching patches exist, in this work an examplar-based inpainting using sparse representation principles and dictionary learning algorithms is used in the experiments.

2.1.1. Sparse Representation

Sparse representations of signals are meaningful under the assumption that the signal can be represented sparsely in a domain, usually represented by atoms collected as columns in a dictionary matrix. The dictionary D is a matrix $N \times K$, which contains K prototype signals of length N, also referred to as atoms. The

model assumes that for any signal x, there exists a sparse linear combination of atoms from D that approximates it well. The approximation of x can be written as

$$x \approx Dw, \qquad \|w\|_0 \ll N \tag{1}$$

where w is a vector containing the coefficients and most of the entries in w are zero and the operator $\|.\|_0$ counts the number of non-zero elements in a vector. Typically it is assumed that the dictionary is redundant in describing x.

So, given the dictionary D, the approximation $\hat{x}$ of signal x can be written as $\hat{x} = Dw$, and the representation error or residue can be written as $r = x - \hat{x} = x - Dw$. Most of the entries of w are zero, s is the number of non-zero coefficients, and s/N is the sparseness factor.

A common way to find w, i.e the sparse approximation problem, is solving the following equation:

$$w_{opt} = \underset{w}{\operatorname{argmin}} \{\|w\|_p + \gamma \|x - Dw\|_2^2\}, \quad p \in \{0, 1\}. \tag{2}$$

The problem with p $= 0$ (to minimize the number of non-zero coefficients) is NP-hard, but an approximate solution can be found by greedy methods. Alternatively the problem can be relaxed by setting p $= 1$ (to minimize the sum of absolute values) providing a convex problem that can be solved (LASSO). Both problem resolutions start with an all zero vector w, which is the solution when γ is close to zero. As γ factor increases the solution is getting more dense.

2.1.2. Dictionary Learning

Dictionary Learning is often formulated as the problem of finding a dictionary such that the approximations of many vectors, the training set, are as good as possible given a sparseness criterion on the coefficients, i.e. allowing only a small number of non-zero coefficients for each approximation. Let X be a matrix containing K signals to be represented. The dictionary learning problem can be formulated as an optimization problem with respect to the coefficient matrix W and the dictionary D as:

6

$$\{D_{opt}, W_{opt}\} = \underset{D,W}{\operatorname{argmin}} \sum_{i=1}^{K} \|w_i\|_p + \gamma \|X - DW\|^2 \qquad (3)$$

Different methods have been proposed to solve the previous optimization approaches. In this work the Recursive Least Squares Dictionary Learning Algorithm (RLS-DLA) [18] was employed. In this method a single training vector x_i or a mini-batch (subset) of X is processed in each iteration solving the equation:

$$D_i = B_i A_i^{-1}, \qquad being \quad \begin{cases} A_i = \lambda_i A_{i-1} + W_i W_i^T \\ B_i = \lambda_i B_{i-1} + X_i W_i^T \end{cases} \qquad (4)$$

where $A_1 = W_1 W_1^T$ and $B_1 = X_1 W_1^T$.

The current dictionary D_{i-1} is used to find the corresponding coefficients W_i. The main improvement in RLS-DLA, compared to Least Square Dictionary Learning Algorithm (LS-DLA)[19], is that instead of calculating the least squares solution in each step, the matrix inversion lemma (Woodbury matrix identity) can be used to update $C_i = A_i^{-1}$ and D_i directly without using A_i and B_i. The advantage of RLS-DLA compared with Method of Optimized Directions (MOD)[20] and K-Singular Value Decomposition (K-SVD)[21] comes with the flexibility introduced by a forgetting factor λ. The idea is to forget more quickly in the beginning and then forget less as learning proceeds and we get more confidence in the quality of the dictionary. This can be done by starting with λ slightly smaller than one and slowly increasing λ towards 1 as learning progresses. The update scheme and λ should be chosen so that the initial dictionary is soon forgotten and convergence is obtained in a reasonable amount of iterations.

2.1.3. Inpainting using Sparse Representation and Dictionary Learning

Sparse representations with learned dictionaries are, among other things, capable of producing state of the art inpainting results [22]. Let x be a length(N) data vector, for example originated from an image patch (squared or from another neighbourhood), with the pixels stacked as a vector. A dictionary de-

7

scribed by a $N \times K$ matrix: $D = [d_1 d_2 \cdots d_K]$ being K the number of background patches from the image to be inpainted (only taking into account the blocks without missing pixels). Assume there exist a dictionary that represents the data well and it is computed using Dictionary Lerning techniques. This optimized dictionary is composed by the L most representative atoms, being $L = 2N$ (Figure 2a).

In inpainting there are missing pixels, and the position of the missing pixels are known. Let us define $\tilde{x}$ of length $(N - p)$ where the p missing pixels are removed. The inpainting process (Figure 2b) consists of:

1. Removing the corresponding rows of D, giving $\tilde{D}$ of size $(N - p) \times K$.

2. Finding w as $min_w||\tilde{x} - \tilde{D}w||$ $s.t.||w||_0 < s$.

3. Finding the reconstructed vector $x_r = Dw$ using the full dictionary.

4. Finally, the missing pixels in $\tilde{x}$ can be replaced by the values in the corresponding positions of x_r, alternatively all $\tilde{x}$ is replaced by x_r if a denoising of all pixels is needed in addition to filling in the missing pixels.
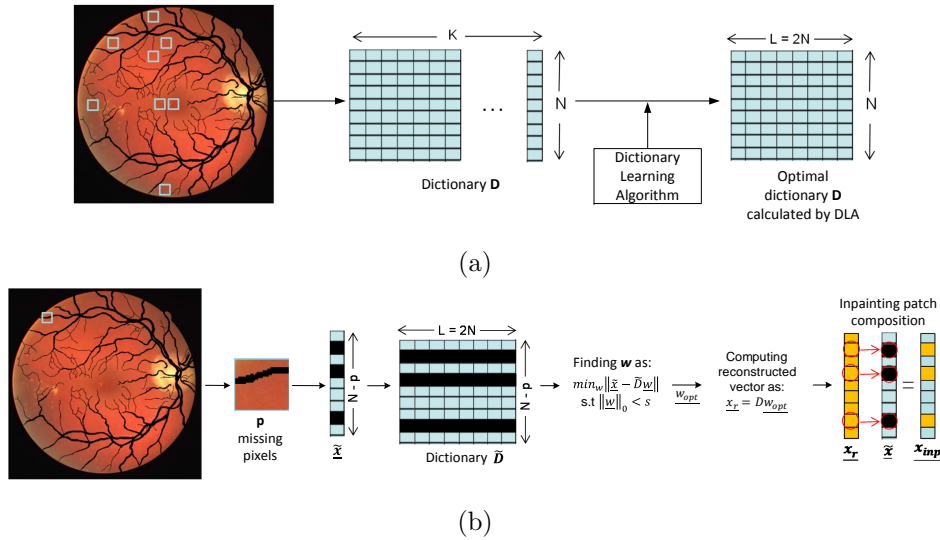

(a)

(b)

Figure 2: Retinal vessel inpainting using Sparse Representation and Dictionary Learning. (a) Dictionary Learning process, (b) Image inpainting process.

As can be observed in Figure 2b, the reconstructed vector of each patch,

8

extracted from the image to be inpainted, is computed. The inpainted patch is composed replacing the missing pixels in the original patch by the values in the corresponding positions of the reconstructed patch. The final inpainted image is formed by the composition between the original image and the inpainted vascular tree (Figure 3).

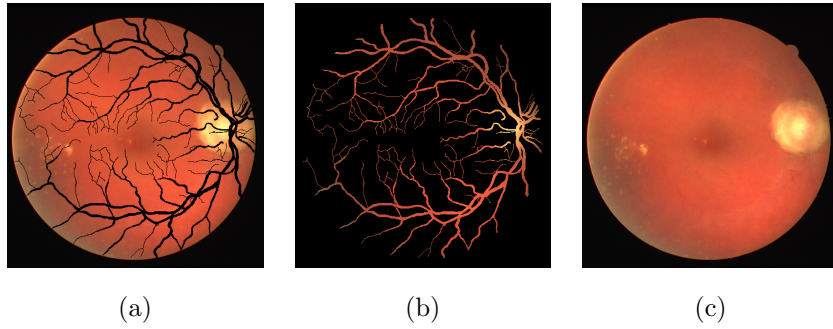(a)   (b)   (c)

Figure 3: (a) Original mask with missing pixels, (b) Inpainting performed on retinal vascular tree (c) Final inpainted image.

Figure 4 shows an overview of the proposed image inpainting method applied to retinal images. In this flow chart it is possible to differentiate the two main stages detailed above: the Dictionary Learning and the image inpainting processes.
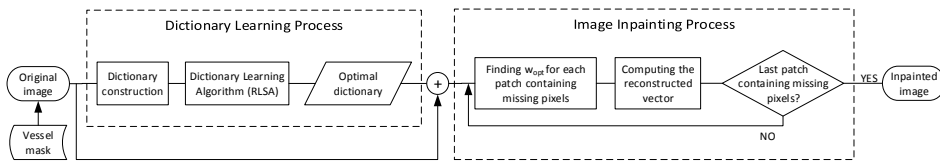

Figure 4: Overview of the proposed image inpainting method applied to retinal images.

The process described above can be implemented defining an image M of the same size as the original image, masking the positions of all the missing pixels and using efficient implementations. In that way, the dictionary matrix $\tilde{D}$ do not have to be explicitly defined. The Orthogonal Matching Pursuit algorithm (with mask) [10] was used in these experiments, but that could easily be changed

9

to a L1 norm algorithm, like LASSO [23], if preferred. SPAMS library, SPArse Modeling Software [24], includes an efficient implementation of this algorithm.

2.2. Inpainting Quality Assessment

The quantification of the quality of the image inpainting offers some challenges. Blood vessel removal is a real application in which the "ideal" fundus image without blood vessels does not exist and thus can't be compared to the inpainted image. In most state of the art work within inpainting the assessment is done by comparing the resulting image to a truth image, while doing inpainting on a corrupted version of the truth image. Two different ways to carry out the inpainting quality assessment are presented in the following and they are used to evaluate the results in each experiment of this work.

2.2.1. Recovery Error

A simple method to measure the image quality inpainting is modifying the vessel mask segmented by the specialists. In particular additional missing pixels (artificial mask) are added to this hand-segmented mask giving place to the missing vessel mask that will be used in the experiments (Figure 5).

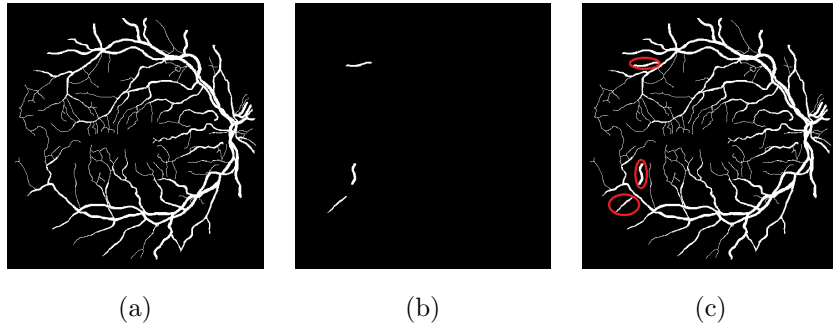(a)   (b)   (c)

Figure 5: An example of mask transformation. (a) Original mask, (b) Artificial mask (c) Missing pixel mask.

Using the Recovery Error (RE) metric it is possible to quantify the quality of the inpainted images. The intensity value of the missing pixels artificially added to the vessel mask is known in the original image, for this reason, it is

10

possible to extract the error due to the inpainting process by computing the difference between the original pixel and the inpainted one at each additional missing pixel. The average of Recovery Error (RE) along the whole additional missing pixel mask provides a good metric of how good the image inpainting is:

$$\text{RE} = \sqrt{\frac{\sum_{n=1}^{N_c} \sum_{k=1}^{M} (Iout(x_k, y_k, n) - Iin(x_k, y_k, n))^2}{M}} \tag{5}$$

where $Iout$ is the resulting image after inpainting and Iin is the original (truth) image, both normalized to the range of 0 to 1 and evaluated at the M pixels (x_k, y_k) of the artificial mask. N_c is the number of components (1 for graylevel images and 3 for RGB). A small RE indicates a high quality inpainting.

2.2.2. Inpainting Quality Evaluation Index

The Recovery Error needs the creation of an artificial mask for each image and also it only measures the quality in the relatively small areas of artificial mask. For these reasons, a novel metric is proposed in order to quantify the quality image inpainting without the use of a reference (truth) image.

We have observed that subjective quality by visual inspection is based on two factors:

- The coherence of the inpainted region and the background, in other words, the low level of correlation between the resulting inpainted region and the same region evaluated at the original image.

- The prominence/smoothness of the resulting edges of the inpainted structure to be removed. The more smooth the edges that joint the inpainted region and the background are, the better the quality of the inpainting.

We propose an automatic Inpainting Quality Evaluation (IQE) index exploiting these two factors, in order to evaluate the resulting images, defined as:

$$\text{IQE} = 1 - [G \times HF \times C] \tag{6}$$

11

where G evaluates the gradient similarity between the original and the inpainted
image, HF measures the ratio of high frequencies between the images and C
quantifies the image correlation. This dissimilarity index is ranged from 0 to 1
and a high IQE value corresponds to a good inpainting quality.

Definition of the gradient similarity and high-frequency terms

Human vision is very sensitive to contrast variations or texture changes in
images. Image edges separate different objects or structures of the image, thus
the gradient can capture well contrast and texture variations well. In addition,
by estimating a high-frequency parameter, it is possible to extract information
about the prominence or smoothness of the contours in the junctions between
the inpainted region and its surrounding background. The G and HF terms focus
on these ideas and both parameters are computed according to the equations
detailed in [17].

Definition of the correlation term

With the aim of quantifying the coherence of an inpainted region and its
neighbourhood a luminance term is proposed in [17]. Using this term, only the
mean intensities between a particular patch from the inpainted image and the
original one can be compared. With this idea, both, pixels with and without
inpainting inside the window contributes in the same way to the final value of
the metric for the patch. Moreover, in the global metric, both, patches with
missing pixels (which suffer inpainting) and patches with no missing pixels,
are taken into account modifying the global value of the metric (increasing the
value, because the patches non suffering inpainting have the maximum value and
contribute overvaluing the metric). For these reasons, we propose a new term
to quantify how coherent an inpainted patch is with the background, or how
well blood vessels have been transformed into retinal texture. For these reasons
we propose a new term to quantify how coherent a new inpainted patch is with
the background, or how well blood vessels have been transformed into retinal
texture. The correlation level between the inpainted image and the original one

12

is computed in each RGB channel. A sliding window, of dimensions $m \times n$, loops the vessel mask. For each window in which at least one pixel is masked as blood vessel, the correlation term $C(x, y)$ is calculated as:

$$C(\mathrm{X, Y}) = \frac{\sum_{i=0}^{m-1} \sum_{j=0}^{n}[X(i,j) - \bar{X}][Y(i,j) - \bar{Y}]}{\sqrt{(\sum_{i=0}^{m-1} \sum_{j=0}^{n-1}[X(i,j) - \bar{X}]^2)(\sum_{i=0}^{m-1} \sum_{j=0}^{n-1}[Y(i,j) - \bar{Y}]^2)}} \quad (7)$$

where $X(i,j)$ is the ith, j^{th} pixel in the sliding window evaluated in the inpainted image and $Y(i,j)$ is the ith, j^{th} pixel in the sliding window evaluated in the original one (image with blood vessels). $\bar{X}$ and $\bar{Y}$ are the average values of the intensity calculated in the inpainted and the original image respectively.

The correlation term $C(A,B)$ between the whole inpainted image A and the original one B is calculated as the average of $C(X,Y)$ terms for each window in which at least one pixel belongs to a blood vessel:

$$C(\mathrm{A, B}) = \frac{1}{N_w} \sum_{k=0}^{N_w - 1} C_k(X, Y) \quad (8)$$

where N_w is the number of windows in which at least one pixel is labelled as blood vessel. A low correlation corresponds to a high coherence between the inpainted region and its surrounding background.

3. Results

Three experiments were carried out. The objective of the first experiment was to assess the influence of the sparseness factor and the neighbourhood. For that the sparse-based method with RLS-DLA on graylevel images were used. In the second experiment, two methods of dictionary learning were used with the aim of measuring the effect of inpainting each component of a RGB image (separate inpainting) versus to inpaint the three components together (jointly inpainting). Finally, the sparse-based inpainting is compared with a diffusion-based algorithm on RGB images.

13

3.1. Material

The material of this work are publicly available fundus images. Two different databases were used. DRIVE [25] database is composed by 40 retinal images (565 x 584 pixels) all belonging to diabetic subjects. For each image, a mask image that delineates the field of view is provided (external mask) as well as manual segmentations of the blood vessels (vessel mask). STARE [26] database is a set of 20 images (700 x 605 pixels) along with two hand labelled vessel network provided by different experts. Note that a list with the original images used as well as the resulting images of the inpainting are made available at http://cvblab.synology.me/MOSevaluation/MOSevaluation.rar to facilitate future fair comparisons and to allow the evaluation of experts in this domain.

3.2. Gray Level Experiment

In this part of the work, the green channel of a RGB image was chosen to perform the vessel inpainting because it is the component commonly used to detect the lesions (the objective of a computer-aided diagnosis system) [27–29].

The original image is divided in L overlapping patches of size $N = BB_{height} \times BB_{width}$ or other non-squared neighbourhoods. The initial dictionary D is composed by atoms from the training signal, in this context that means that atoms are constructed by means of normalized stacked pixels of patches (without any missing pixels) extracted from the image. The RLS-DLA [18] is applied in order to optimize the dictionary D, (size $N \times K$). The new dictionary is composed by the L most relevant atoms being $L = 2N$, so the learned dictionary is a matrix $N \times L$, which contains L prototype signals of length N.

Inpainted images are obtained using the learned dictionary and the fast implementation of the OMP method (with mask) provided in the SPAMS library [24].

Different neighbourhoods were tested: squared blocks of sizes 8×8, 21×3 and 3×21. The choice of the neighbourhood size is a key decision because is directly related with the size of the anatomical structure to be inpainted and, therefore, with the image resolution. In this work, two public databases are used

14

and the different resolutions affects in the neighbourhood size decision. A spatial resizing was performed in the STARE images in order to set the same resolution as the DRIVE images. Using these image dimensions (565 x 584 pixels), the thickest blood vessel takes up around five pixels. So, the patch size (in each dimension) should be higher than the size of the biggest vessel, to avoid patches containing only missing pixels. Finally, the patch size is chosen as a trade-off between the sparsity of zeros inside the patch and the computational cost. In this work square and rectangular patches as well as sparser neighbourhoods were tested in order to compare the behavior.

The inpainting based on sparse representation with squared neighbourhoods works well when the missing pixel areas are small [9]. In our problem, we have continuous, and large, missing pixels areas which are not properly addressed by this kind of neighbourhoods. Our idea is to use more spread neighbourhoods trying to minimize the number of missing pixels in the same patch. So, in addition to the squared neighbourhoods, the patterns shown in Figure 6 were tested. These patterns contain 77, 43 and 115 elements respectively. The same neighbourhood is used for learning the dictionary as used later in the inpainting step.

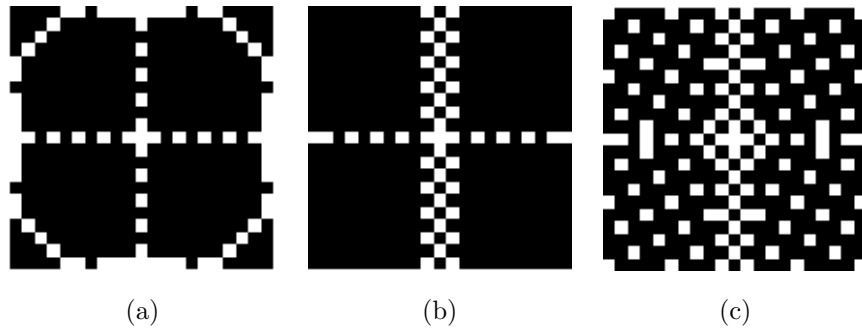(a)   (b)   (c)

Figure 6: Tested neighbourhoods. (a) p1, (b) p2 (c) p3.

Table 1 shows the results of the average recovery error (for the forty images from the DRIVE database) in the different tested configurations (8×8, 21×3, 3×21, p1, p2 and p3 neighbourhoods). The tested number of non-zero

15

coefficients of vector w were s=4 and s=10.

	8x8	21x3	3x21	p1	p2	p3
s=4	0.0386	0.0409	0.0431	0.0286	**0.0275**	0.0291
s=10	0.0390	0.0414	0.0432	0.0283	**0.0272**	0.0290

Table 1: Mean Recovery Error calculated from the resulting images of grayscale inpainting using RLS-DLA for different values of sparseness factors (s=4 and s=10) and different neighbourhoods.

As can be observed in Table 1, results are slightly worse in the case of squared neighbourhoods. However, in relation to the non-zero coefficients (sparseness factor) the performance is similar using $s = 4$ and $s = 10$.

Figure 7 shows an original image and the inpainting result for the best case (neighbourhood p2).
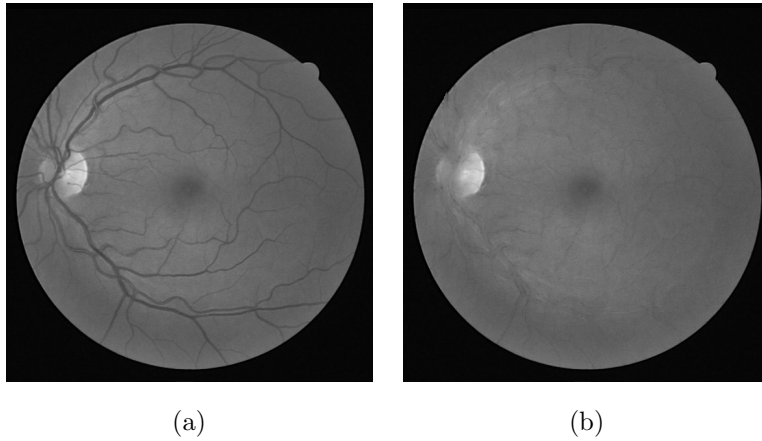

(a)   (b)

Figure 7: Grayscale inpainting for the 01_test image from DRIVE public database. (a)Original image and (c)Inpainted image. This result was obtained using RLS-DLA with the sparseness factor s = 10 and the second spread neighbourhood.

3.3. RGB Experiment

This experiment aims to analyse which configuration provides better results: to inpaint each color component separately or the jointly inpainting of the three

components [22]. Another objective in this experiment is to demonstrate if the choice of the dictionary learning algorithm (in order to create the optimal dictionary) influences in the quality of the resulting inpainted image. DRIVE database and recovery error were used in this experiment.

3.3.1. RGB-separate Inpainting

Each component of RGB images were extracted and inpainted following the same method as presented in the graylevel experiment. The resulting RGB image is the composition of the three inpainted images, one for each channel. In this experiment the performance of the inpainting employing the RLS-DLA is compared with the performance of the result using Online Dictionary Learning (ODL) algorithm [30].

Table 2 shows the results of the mean recovery error (for the whole DRIVE dataset) measured on the composed RGB image, for RLS-DL and ODL algorithms.

		8x8	21x3	3x21	p1	p2	p3
RLS-DL	s=4	0.0578	0.0549	0.0651	0.0496	0.0506	**0.0463**
	s=10	0.0575	0.0553	0.0645	0.0489	0.0503	**0.0453**
ODL	s=4	0.0559	0.0577	0.0662	0.0474	0.0518	0.0471
	s=10	0.0567	0.0581	0.0641	0.0473	0.0512	0.0464

Table 2: Mean Recovery Error calculated from the resulting images of RGB-separate inpainting using RLS-DL and ODL algorithms for different values of sparseness factors (s=4 and s=10) and different neighbourhoods.

3.3.2. RGB-jointly Inpainting

In this method an atom of the dictionary is composed by the same patch in each channel, so the learned dictionary D' is a matrix $N' \times L'$ where $N' = 3\ N$ and $L' = 2\ N'$. In this case the number of non-zero coefficients of vector w' should be tripled, compared to the gray level experiment, because its length is three times the length of w. The inpainted RGB image is obtained in one step and the recovery error is measured on this image directly.

17

Table 3 shows the results of the mean recovery error for the RGB image using RLS-DL and ODL algorithms.

		8x8	21x3	3x21	p1	p2	p3
RLS-DL	s=12	0.0377	0.0396	0.0496	**0.0225**	0.0253	0.0283
	s=30	0.0382	0.0396	0.0499	**0.0213**	0.0249	0.0281
ODL	s=12	0.0382	0.0396	0.0470	0.0249	0.0274	0.0272
	s=30	0.0387	0.0384	0.0486	0.0242	0.0270	0.0280

Table 3: Mean Recovery Error calculated from the resulting images of RGB-jointly inpainting using RLS-DL and ODL algorithms for different values of sparseness factors (s=12 and s=30) and different neighbourhoods.

Comparing Table 2 and Table 3 an improvement for all neighbourhoods is observed inpainting the three channels in one step, rather then individually. Thus, it is beneficial using one dictionary in which each atom is composed for a patch extracted from all channels.

Figure 8 displays the results of two inpainted images using the proposed method (with the best result, i.e with p1 and RLS-DLA). Figure 8a shows an image that includes exudates and Figure 8c shows an image with microaneurisms and haemorrhages. As can be observed from the output images, (Figure 8b and Figure 8d), the blood vessels are significantly suppressed without any visual degrading of the exudates or microaneurisms.

3.4. Comparison of Inpainting Methods

With the aim of characterizing the best inpainting category for our application, a comparison between diffusion-based inpainting, the inpainting method used in the retina by other authors, and sparse-based inpainting, the proposed method in this paper, was carried out and the results are presented in this section.

Besides DRIVE database, in order to provide a wider comparison, the fundus images of the STARE public database, introduced in section 3.1, were added in this experiment. In order to avoid the task of the artificial mask extraction for
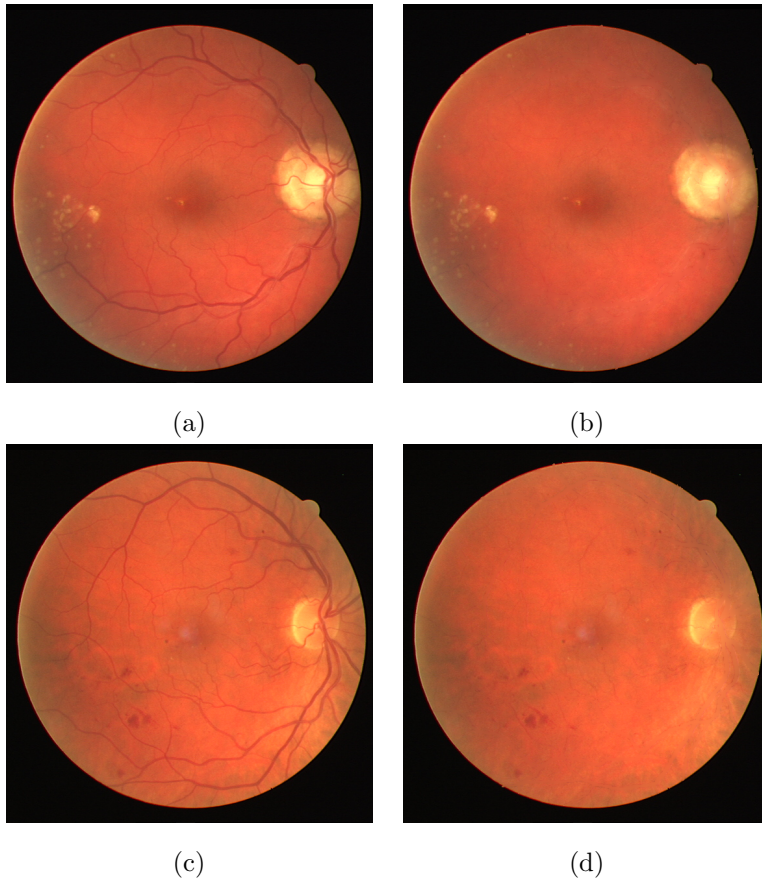
Figure 8: RGB-jointly inpainting for two images. Original images (a)08_test and (c)14_test from DRIVE public database. Inpainted images (b)08_test (RE = 0.207) and (d)14_test (RE = 0.211). The inpainting was performed using RLS-DLA with the sparseness factor s = 30 and the first spread neighbourhood.

each image belonging to STARE database, the proposed IQE index was used to quantify the quality of the blood vessel inpainting.

Diffusion-based inpainting category by means of the simplest isotropic diffu-
385 sion model [31] and a public implementation of the exemplar-based inpainting method presented by Criminisi et al.[32] were selected to perform the comparison with the proposed sparse-based inpainting method (Table 4). The same neighbourhoods used in the previous experiments were tested for the sparse-

based method.

In order to validate the IQE index, the RGB-jointly inpainting and the RGB-separate inpainting were carried out. Note that in relation to the non-zero coefficients, $s = 10$ was used for RGB-separate (RGBS) inpainting and $s = 30$ for the RGB-jointly (RGBJ) inpainting. In addition, a Mean Opinion Score (MOS) evaluation of the inpainting results was performed by five experts in digital image processing (Table 5). These experts assessed the 60 resulting images according to the following quality scale: Very poor(0), Poor(1), Fair(2), Good(3), and Excellent(4). In the evaluation process, the resulting images were presented in random order to the experts and without any kind of information about the inpainting method used to remove the vessels.

		8x8	21x3	3x21	p1	p2	p3	DBI	[32]
DRIVE	RGBS	0.541	0.5276	0.5189	0.6044	0.597	0.617	0.5039	0.6641
	RGBJ	0.5596	0.5461	0.535	0.6838	0.6194	**0.6915**		
STARE	RGBS	0.6224	0.6013	0.5897	0.6539	0.6562	0.6524	0.5748	**0.6691**
	RGBJ	0.6152	0.6094	0.5979	**0.6648**	0.6572	0.6582		

Table 4: IQE index average for the RGB-jointly inpainting (s = 30) and the RGB-separate inpainting (s = 10) computed from DRIVE and STARE databases using diffusion-based inpainting (DBI), Criminisi et al. [32] method and different neighbourhood configurations of sparse-based Inpainting.

		8x8	21x3	3x21	p1	p2	p3	DBI	[32]
DRIVE	RGBS	0.3	0.05	0.075	3.35	1.575	3.625	1.625	2.75
	RGBJ	0.65	0.18	0.22	3.475	2.275	**3.7**		
STARE	RGBS	0.388	0.117	0.21	2.95	2.1	2.65	1.5	**3.1**
	RGBJ	0.75	0.33	0.412	**2.95**	2.2	2.65		

Table 5: Mean Opinion Score (MOS) evaluation of the inpainting results obtained from the DRIVE and STARE databases. The assessment was carried out using the following scale: Very poor(0), Poor(1), Fair(2), Good(3), and Excellent(4).

The novel IQE index allows to quantify the performance of the inpainting

result taking into account the result in the whole image, not only in a few pixels artificially added composing an artificial mask. For this reason, comparing the results for DRIVE database (Table 4) with the results presented in Tables 2 and 3 different results are observed. In Table 3 the best results are obtained using the pattern one, but assessing the inpainting quality by the IQE index the best pattern is found as number three in Table 4. We believe that the IQE index provides us a more reliable measure of the quality of the inpainting compared to the Recovery Error because RE only evaluates the quality in the relatively pixels of the artificial mask.

As can be observed in Table 4, the IQE index ratifies that the RGB-jointly inpainting of the three channels provides better results than the RGB-separate inpainting. The results of the MOS evaluation (Table 5) shows high correlation with the IQE results, demonstrating the reliability of the proposed metric.

Diffusion-based techniques perform well filling small gaps but the results are quite poor when the unknown region to be inpainted is large and blood vessels are a large structure that covers a high percentage of the fundus image. This fact could be the consequence of the humble results observed in Table 4 when diffusion-based inpainting is used.

The method proposed by Criminisi et al.[32] is based on the idea of finding the image patch that maximises the correlation with the region to be inpainted. Table 4 shows similar results using this method in comparison with the best configuration of the proposed sparse-based method. Figure 9 allows a visual comparison of the resulting images using the inpainting methods compared in this work.

Despite the well performance using the Criminisi et al.[32] method, the public implementation used in this work is really time-consuming. In order to check this fact, Table 6 shows the average time taken by the different algorithms compared in this work when blood vessel removal is performed in DRIVE and STARE databases.

Note that comparison with other work where inpainting is done on fundus images is not possible because the inpainting techniques are used as a middle
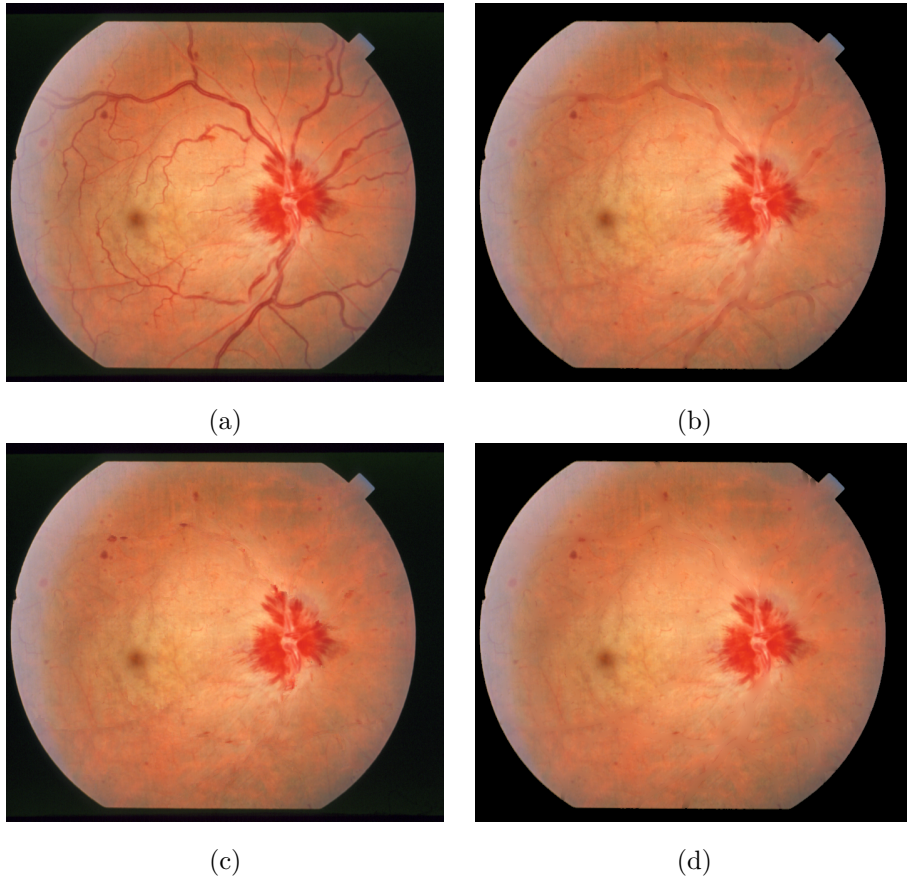
21

Figure 9: im0005 from STARE public database. (a) original, (b) Inpainted image by diffusion-based method (IQE = 0.5322), (c) Inpainted image by Criminisi et al. method [32] (IQE = 0.6631) and (d) RGB-jointly inpainted (s = 30) image by sparse-based method using the first pattern (IQE = 0.6886).

step, and the inpainting quality is not quantitatively validated, to the authors knowledge.

4. Conclusions

In this paper, different inpainting techniques were applied to remove blood vessels of the fundus images. We propose to use a sparse-based method for inpainting this application, and experiment on RGB-jointly versus RGB-seperate

22

	Sparse-based	Diffusion-based	Criminisi et al.[32]
DRIVE	297.07 sec.	91.78 sec.	609.41 sec.
STARE	303.64 sec.	98.33 sec.	647.58 sec.

Table 6: Averaged time (in sec.) consumed for the Sparse-based method, Diffusion-based method and Criminisi et al. method to perform the blood vessel removal in each database involved in this work.

configurations. Different patch patterns as well as sizes are investigated. A diffusion-based method is also tested for reference. Different experiments were carried out using two public databases of fundus images.

Two different metrics were proposed and used in order to quantify the quality of the inpainted images (in non-artificial applications) and compare the different methods and configurations tested. The RE is based on evaluating artificially added missing pixel areas. In this paper we propose the IQE index, which provides more efficient and reliable results than the Recovery Error for two main reasons. The IQE index does not need the creation of an artificial mask. In addition, this metric allows to quantify the performance of the inpainting result taking into account the whole image, not only a small area corresponding to the artificial mask.

Through dictionary learning algorithms, the optimal dictionaries from DRIVE database were obtained for different patterns, and the pixels belonging to the vessels were inpainted by combining the pixels of the patches that produce the optimal coefficients w. Results show a better performance using spread patterns rather than using square neighbourhoods.

Two different ways to obtain the inpainted RGB image have been tested. Learning the dictionary taking into account the three channels provided better results than to inpaint each component separately.

Moreover, RLS-DL technique was compared with ODL obtaining comparable results, demonstrating that inpainting techniques based on sparse representation and dictionary learning can be used to remove vessels in retinal images.

23

Regarding the sparseness factor, results show that this factor has not a high influence in the result.

Blood vessels are a large structure that covers a high percentage of the fundus images and that might be part of the reason why sparse-based methods perform better than the diffusion based. The proposed method has also been compared with the exemplar-based inpainting method by Criminisi et al. [32] obtaining similar results in a more efficient way.

It is possible to obtain an image better suited for texture analysis and lesion segmentation by firstly perform sparse based inpainting of the RGB-jointly image using a spread patch mask. This can hopefully lead to a more reliable classification between pathological and healthy texture and a more dependable diagnosis of retinal diseases.

In future work the proposed method using sparse-based algorithms and the proper configuration will be used in order to obtain a retinal image without blood vessels. The final objective of the project is to develop an automatic screening software in order to distinguish between the normal and pathological retina and classify these pathologies. The previous step to the inpainting stage is to segment automatically the retinal blood vessels. In a previous work, the authors proposed an automatic algorithm for this purpose [33]. Next step to the inpainting stage is the texture characterization [34], using as input the inpainted images.

Acknowledgments

24

References

[1] World Health Organization (WHO), Action plan for the prevention of avoidable blindness and visual impairment 2009-2013, Tech. rep. (2010).

[2] C. Guillemot, O. Le Meur, Image inpainting: Overview and recent advances, IEEE Signal Processing Magazine 31 (1) (2014) 127–144.

[3] M. Bertalmio, A. L. Bertozzi, G. Sapiro, Navier-stokes, fluid dynamics, and image and video inpainting, in: Proc. IEEE Computer Vision and Pattern Recognition (CVPR), 2001, pp. 355–362.

[4] A. Telea, An image inpainting technique based on the fast marching method., J. Graphics, GPU, & Game Tools 9 (1) (2004) 23–34.

[5] D. Tschumperl, Fast anisotropic smoothing of multi-valued images using curvature-preserving pde's, International Journal of Computer Vision 68 (1) (2006) 65–82.

[6] L. Liang, C. Liu, Y.-Q. Xu, B. Guo, H.-Y. Shum, Real-time texture synthesis by patch-based sampling, ACM Trans. Graph. 20 (3) (2001) 127–150.

[7] D. Doria, R. J. Radke, Filling large holes in lidar data by inpainting depth gradients, in: CVPR Workshops, IEEE, 2012, pp. 65–72.

[8] H.-Y. Huang, C.-N. Hsiao, A patch-based image inpainting based on structure consistence, in: Computer Symposium (ICS), 2010 International, IEEE, 2010, pp. 165–170.

[9] J. Mairal, F. Bach, J. Ponce, G. Sapiro, A. Zisserman, Discriminative learned dictionaries for local image analysis., in: CVPR, IEEE Computer Society, 2008.

[10] M. Elad, Sparse and redundant representations: from theory to applications in signal and image processing, Springer, 2010.

[11] T. T. Dang, A. Beghdadi, M.-C. Larabi, Perceptual quality assessment for color image inpainting, in: IEEE International Conference on Image Processing, 2013, pp. 398–402.

[12] T. T. Dang, A. Beghdadi, M.-C. Larabi, Visual coherence metric for evaluation of color image restoration, in: IEEE Colour and Visual Computing Symposium, 2013, pp. 1–6.

[13] V. Voronin, V. Marchuk, E. Semenishchev, S. Maslennikov, I. Svirin, Inpainted image quality assessment based on machine learning, Conference on Computer Graphics, Visualization and Computer Vision, 2015.

[14] A. Giachetti, K. Chin, E. Trucco, C. Cobb, P. Wilson, Multiresolution localization and segmentation of the optical disc in fundus images using inpainted background and vessel information, Image Processing, IEEE International Conference on (2011) 2145–2148.

[15] S. Morales, V. Naranjo, J. Angulo, M. Alcaniz, Automatic detection of optic disc based on pca and mathematical morphology, Medical Imaging, IEEE Transactions on 32 (4) (2013) 786–796.

[16] J. Meier, R. Bock, G. Michelson, L. Nyúl, J. Hornegger, Effects of preprocessing eye fundus images on appearance based glaucoma classification, in: Proceedings of the 12th International Conference on Computer Analysis of Images and Patterns, Springer-Verlag, 2007, pp. 165–172.

[17] S. Wang, H. Li, X. Zhu, P. Li, An evaluation index based on parameter weight for image inpainting quality, in: Young Computer Scientists, 2008. The 9th International Conference for, IEEE, 2008, pp. 786–790.

[18] K. Skretting, K. Engan, Recursive least squares dictionary learning algorithm, Signal Processing, IEEE Transactions on 58 (2010) 2121–2130.

[19] K. Engan, K. Skretting, J. Hakon Husy, Family of iterative ls-based dictionary learning algorithms, ils-dla, for sparse signal representation, Digital Signal Processing 17 (1) (2007) 32 – 49.

26

[20] K. Engan, S. O. Aase, J. Hakon Husoy, Method of optimal directions for frame design, in: Proceedings of the Acoustics, Speech, and Signal Processing, 1999. On 1999 IEEE International Conference - Volume 05, ICASSP '99, IEEE Computer Society, 1999, pp. 2443–2446.

[21] M. Aharon, M. Elad, A. Bruckstein, The k-svd: An algorithm for designing overcomplete dictionaries for sparse representation, Trans. Sig. Proc. 54 (11) (2006) 4311–4322.

[22] J. Mairal, M. Elad, G. Sapiro, Sparse representation for color image restoration, Image Processing, IEEE Transactions on 17 (1) (2008) 53–69.

[23] R. Tibshirani, Regression shrinkage and selection via the lasso, Journal of the Royal Statistical Society. Series B (Methodological) (1996) 267–288.

[24] J. Mairal, F. Bach, J. Ponce, G. Sapiro, R. Jenatton, G. Obozinski, Spams v2.5 (sparse modeling software), `http://spams-devel.gforge.inria.fr/index.html` (2014).

[25] J. Staal, M. Abrmoff, M. Niemeijer, M. Viergever, B. van Ginneken, Ridge-based vessel segmentation in color images of the retina., IEEE Trans. Med. Imaging 23 (4) (2004) 501–509.

[26] A. Hoover, V. Kouznetsova, M. Goldbaum, Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response, IEEE Transactions on Medical Imaging 19 (2000) 203–210.

[27] X. Zhang, G. Thibault, E. Decencire, B. Marcotegui, et al., Exudate detection in color retinal images for mass screening of diabetic retinopathy, Medical Image Analysis 18 (7) (2014) 1026–1043.

[28] M. Abràmoff, J. Folk, D. Han, et al., Automated analysis of retinal images for detection of referable diabetic retinopathy, JAMA Ophthalmology 131 (3) (2013) 351–357.

[29] T. Walter, J.-C. Klein, P. Massin, A. Erginay, A contribution of image processing to the diagnosis of diabetic retinopathy - detection of exudates in color fundus images of the human retina., IEEE Trans. Med. Imaging 21 (10) (2002) 1236–1243.

[30] J. Mairal, F. Bach, J. Ponce, G. Sapiro, Online dictionary learning for sparse coding, in: Proceedings of the 26th Annual International Conference on Machine Learning, ACM, 2009, pp. 689–696.

[31] J. D'Errico, Inpainting nans, `http://www.mathworks.com/matlabcentral/fileexchange/4551-inpaint-nans`. Last accessed on 12th January 2016 (2004).

[32] A. Criminisi, P. Prez, K. Toyama, Region filling and object removal by exemplar-based image inpainting, IEEE Transactions on Image Processing 13 (2004) 1200–1212.

[33] S. Morales, V. Naranjo, J. Angulo, J. J. Fuertes, M. Alcañiz, Segmentation and analysis of retinal vascular tree from fundus images processing., in: BIOSIGNALS, SciTePress, 2012, pp. 321–324.

[34] S. Morales, K. Engan, V. Naranjo, A. Colomer, Retinal disease screening through local binary patterns., IEEE Journal of Biomedical and Healths Informatics (Article in press, 2016).

28