

An Efficient Approach for Detecting Exemplar based Image Inpainting and Copy Move Forgery

Rinia Mary Varghese

PG Student

Department of Computer Science & Engineering

Federal Institute of Science and Technology

Abstract

Today manipulation of digital images has become easy due to powerful computers, advanced photo-editing software packages and high resolution capturing devices. Verifying the integrity of images and detecting traces of tampering without requiring extra prior knowledge of the image content or any embedded watermarks is an important research field. Exemplar based inpainting and copy paste forgeries play visual tricks to deceive people and affect the authenticity of the images. The proposed method presents an efficient forgery detection algorithm for detecting the manipulation of images by exemplar inpainting and copy move forgery. The method shows an algorithm for object removal by exemplar-based image inpainting mechanism that restores the area of the removal object and the mechanism can be further employed to extract the back ground of images and it also detect the copy move forgery. Exemplar-based inpainting methods iteratively search the source region and fill the missing or damaged region, i.e., target region, with the most similar patch in the source region. The new method also shows a technique for detecting copy-move forgery over images tampered by copy-move. The proposed method integrates central pixel mapping (CPM), greatest zer connectivity component labeling (GZCL) and fragment splicing detection (FSD). CPM speeds up block search by efficiently matching those blocks with similar hash values and then finding the suspicious pairs. To improve the detection precision, GZCL is used to mark the tampered pixels in suspected block pairs. FSD is adopted t distinguish and locate tampered regions from its best-match regions. The algorithm can reduce the processing time and maintain a detection precision under different kinds of object-removed images. Detecting two forgeries by a single method makes it unique.

Keywords: Image Inpainting, Forgery Detection, CPM

I. INTRODUCTION

For digital images, the manipulations of image objects including object adding, removing or modifying are of the most attention because these changes of objects will directly mislead the understanding and awareness of the image content. In general, image object removal techniques can be grouped into two categories [1]: copy-move and image inpainting. Copy-move is achieved by copying a region from an image and then pasting it into the same image with the intent of hiding undesired objects [2]. Due to its simplicity, copy-move has become the most widely used method for manipulating the semantics of an image. Image Inpainting is the art of filling in missing data in an image. The purpose of inpainting is to reconstruct missing regions in a visually plausible manner so that it seems reasonable to the human eye.

The two forgeries highlighted in this proposed method namely, image inpainting and copy move forgery. Image inpainting is a technique used for reconstruction or image restoration that can recover the lost information in old photographs and remove scratches in images. The term inpainting comes from art restoration, where it is also called retouching In photography and cinema, is used for film restoration, to reverse the deterioration, example cracks in photograph and scratches and dust spots in film. And it is also used for red eye correction, removing object to creative effect and removing stamped date from photographs. The technique can be used to replace the lost blocks in the coding and transmission of images.

Exemplar-based image inpainting is used for removing large objects from an image and fill that whole area with suitable background contents. Image inpainting fills the missing or damaged region in an image utilizing spatial information of neighboring region.

In the proposed work, an efficient forgery detection method is proposed for detecting exemplar image inpainting and copy move forgery. Firstly, the object removal can be done using central pixel mapping (CPM), greatest zero-connectivity component labeling (GZCL) and fragment splicing detection (FSD). The method follows: Firstly, CPM is proposed to speed up block search, which assigns hash value for target block according to the color information of central pixel and then searches for the best-match block among those with similar hash values. The hash values obtained from CPM can better represent the color distributions of image blocks, producing a higher aggregation of similar blocks, which further narrowing the search range. Meanwhile, CPM generates one to one mapping between image blocks and hash values, so that the load factor of the hash table is more close to 1, thus the search efficiency is further improved. Above all, CPM enhances search performance in terms of both load factor and search range. Secondly, GZCL is adopted to mark the tampered pixels in suspicious blocks. After searching the suspicious blocks, the existing methods directly label all the pixels in suspicious block as being tampered, which will lead to higher false alarm and a jagged edge of the detected region. In order to improve detection precision, GZCL takes advantage of

the greatest zero connectivity component. Thirdly, FSD is used to distinguish tampered regions from reference regions and then a histogram is created to obtain the final position of forgery. Most existin algorithms consider both reference and tampered regions as being forged, which makes the reference regions to be another reason of false alarm. In order to solve this problem, FSD calculates the number of regions matched with the target region and filters out those with less matched regions, resulting in a higher detection precision. After that the input image is considered to the corresponding category, whether it is belongs to copy move or exemplar inpainting.

II. RELATED WORKS

There are many existing techniques proposed by previous researchers for detecting image forgeries. In [3] H. T. Sencar et al proposed a new approach for detecting copy-move forgery in digital images, which is considerably more robust to lossy compression, scaling and rotation type of manipulations. Copy-move forgery is a specific type of image tampering, where a part of the image is copied and pasted on another part of the same image. This method used to improve the computational complexity in detecting the duplicated image regions, they propose to use the notion of counting bloom filters as an alternative to lexicographic sorting, which is a common component of most of the proposed copy-move forgery detection schemes. In this work, they propose to extract features from the image blocks by using Fourier-Mellin Transform (FMT). These features would not be only robust to lossy JPEG compression, blurring, or noise addition, but also known to be scaling and translation invariant. Results shows that the proposed features can detect duplicated region in the images very accurately, even when the copied region was undergone severe image manipulations. In addition, it is observed that use of counting bloom filters offers a considerable improvement in time efficiency at the expense of a slight reduction in the robustness.

In [4] C.C. Chang propose a fast and effective forgery detection algorithm for exemplar-based inpainting forged image to recognize tampered inpainting images, which is one of the effective approaches for image manipulation. The proposed algorithm contains two major processes: suspicious region detection and forged region identification. Suspicious region detection searches the similar blocks in an image to find the suspicious regions and uses a similarity vector field to remove the false positives caused by uniform area. It selects candidates and then groups them into several suspicious regions. Forged region identification applies a new method, multi-region relation (MRR) of these suspicious regions and identify the forged regions from the suspicious regions. The proposed approach can effectively recognize if an image is a forged one and identify the forged regions, even for the images containing the uniform background. Moreover, searching the most similar blocks is a time-consuming process instead of an exhaustive search they propose a two-stage searching algorithm based on weight transformation to speed up the computation speed. There are still some limitations in this work. First, an assumption is that the forged region cannot be too small. Therefore, this system is incapable of locating forged regions of small sizes. Second, since the proposed method is based on locating similar parts within an image, they can only deal with the forged techniques such as copy move or inpainting techniques that copy verbatim patches from inside the image. If the forged region is further processed by some techniques like blurring, or the forged region is from different sources, the work may fail.

In [5] G. Yang et. al proposed an efficient forgery detection algorithm for object removal by exemplar-based inpainting, which integrates central pixel mapping (CPM), greatest zero-connectivity component labeling (GZCL) and fragment splicing detection (FSD). CPM speeds up suspicious block search by efficiently matching those blocks with similar hash values and then finding the suspicious pairs. To improve the detection precision, GZCL is used to mark the tampered pixels in suspected block pairs. FSD is adopted to distinguish and locate tampered regions from its best-match regions. Experimental results show that the this algorithm can reduce the processing time and maintain a detection precision under different kinds of object-removed images.

III. PROPOSED METHODOLOGY

The main aim of the technique is to detect the forgery which is affected by exemplar-based image inpainting and copy move forgery. The proposed method is based on CPM, GZCL and FSD. The flowchart of the proposed method shows in Fig.1. First of all, a fast searching algorithm is applied which is known as central pixel mapping. It improve search efficiency and also maintain good detection results. Then suspicious block searching feature is used. After that, GZCL is used to mark the tampered pixels in suspicious blocks. Finally, FSD is employed to filter out the reference regions and get the final locations of tampered regions.

A. Central Pixel Mapping

Central pixel mapping is a fast searching algorithm. Several methods have been proposed to speed up the searching process [4,6]. In [6], a randomized searching method was used to increase the computational efficiency, but it may fall into the local optimization. Then, Chang et al. [4] presented a weight transformation based searching strategy which reduced the searching time and maintained a relatively better detection accuracy. However, it could not optimize load factor of hash table and search range simultaneously, thereby constrained further improvement of search speed and detection accuracy. In order to enhance both load factor and search range, CPM algorithm is used in the new method, which transforms the color information of central pixel in target block into a hash value and search for the best-match block in those blocks with similar hash values.

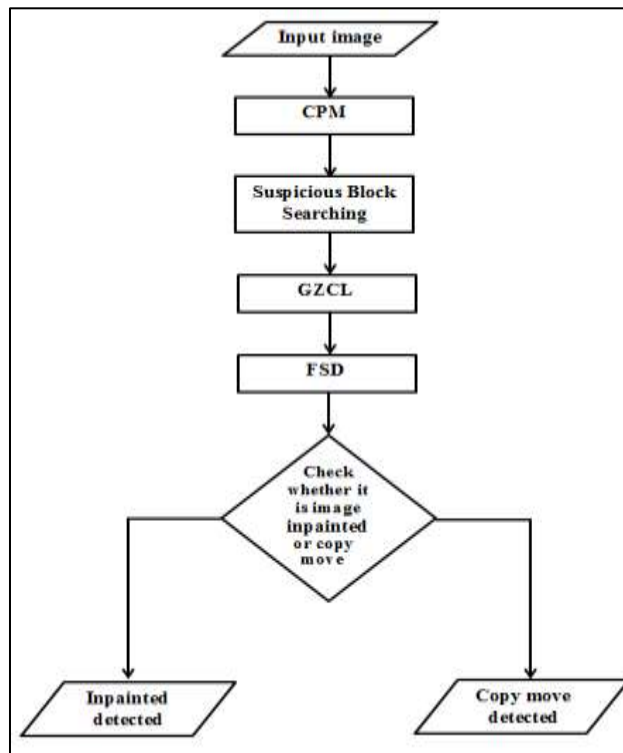


Fig. 1: Flow chart of the proposed technique

1) Block Mapping Technique

The basic idea of block mapping is as follows. Firstly, a hash function is constructed to extract color information of the blocks in the image. In order to speed up the search process, matching degree computation between target block and “alien” blocks, which are entirely different from the target one, should be avoided, and gathering blocks with similar color distribution is necessary. As a result, the hash function also needs to follow this principle, assigning similar hash values to similar image blocks. Secondly, the blocks are mapped to a hash table according to their hash values. However, collision will appear when different blocks correspond to the same hash value. To solve this problem, we store the blocks with same hash value in the same list.

Fig. 2 shows the process of block mapping, where each textural region is represented by a unique color, Value represents the hash value and the dotted red arrows point to the storage location of the blocks. It is not difficult to find that similar blocks are assigned equal or similar hash values. After mapping, the search of suspicious block consists of two steps: (1) locate the target block in the hash table (2) find the best-match block within those with equal or similar hash values. Only the matching degrees among similar blocks are calculated through block mapping, resulting in reduced time complexity and increased search efficiency.

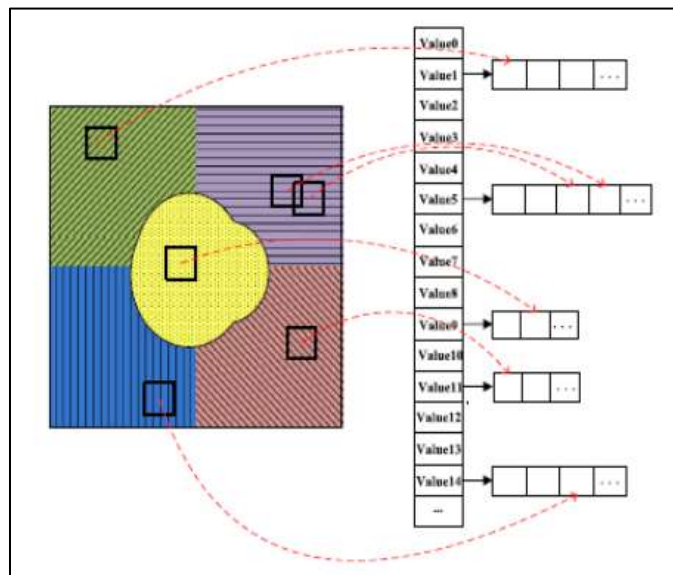


Fig. 2: Block Mapping

2) Principle of Central Pixel Mapping

The key to block mapping is the construction of hash function which will directly affect the search efficiency. In [4], four kinds of hash functions are discussed: equal number matrix transformation, even number matrix transformation, odd number matrix transformation and prime number matrix transformation. Among them, equal number matrix transformation is the best hash function to extract color information from image blocks. However, it tends to cause a collision which enlarges the load factor of the hash table, while the others can deal with collision but fail to gather the similar blocks together, which in turn expands the search range. Therefore, the above mentioned functions cannot optimize load factor and search range simultaneously, which constrains further improvement of search speed. In this work, a new hash function named central pixel mapping (CPM) is defined to gather the blocks with equal value in the central pixel:

$$Key = R(c) + G(c) \cdot 2^8 + B(c) \cdot 2^{16} \quad (1)$$

Where c is the central pixel of target block, R , G , B represent for the values in red, green and blue channel respectively.

According to [7], the central pixel of target block is always located at the edge of the region to be inpainted, which means that there exist pixels to be filled in the 8-neighborhood of the central pixel. Similarly, for forgery detection, when the search window has the same size with the inpainting window, there must be tampered pixels in the 8-neighborhood of central pixel. Moreover, the central pixel is also tampered in the case that the search window is smaller than the inpainting window. According to the filling mechanism of exemplar-based inpainting and copy move, tampered pixels will share the same values with the corresponding ones in the best-match block. ie; the center position of the target block is the same as that of its reference block. Therefore, the search for suspicious blocks can be achieved by only checking the blocks with same central pixels, which verifies that CPM is a proper hash function.

B. Suspicious Block Searching

In exemplar-based inpainting, the unknown part of the target block is filled using the corresponding pixels of reference block and in copy-move forgery, the image is copied and pasted on another part of the same image. Both are resulting in a number of zeros connecting in the difference array between target block and its reference block, that is, a zero is 8 - neighborhood connected with other zeros. This characteristic is named as zero-connectivity feature [8]. Due to its simplicity and effectiveness, zero-connectivity feature is also adopted in this work to search for suspicious blocks.

The matching degree is computed following four steps. Firstly, one block is subtracted from another to get the absolute values of differences for R , G and B components, which are denoted by ΔR , ΔG and ΔB respectively. Then, ΔR , ΔG and ΔB are converted to binary difference arrays D_R , D_G and D_B respectively. Next, mathematical "OR" is performed on D_R , D_G and D_B producing the difference array D_{pq} . At last, the greatest zero-connectivity component which contains the largest number of '0' is taken as the matching degree of the block pair. In order to find the zero-connectivity components in D_{pq} , an 8 - connectedness labeling method is used to extract zero-connectivity features. Fig. 3 illustrates how to compute the matching degree, where R_p , G_p , B_p , R_q , G_q and B_q represent the R , G , B components of target block and reference block respectively, while the zero connectivity components are labeled with green lines. Here, the matching degree is 8.

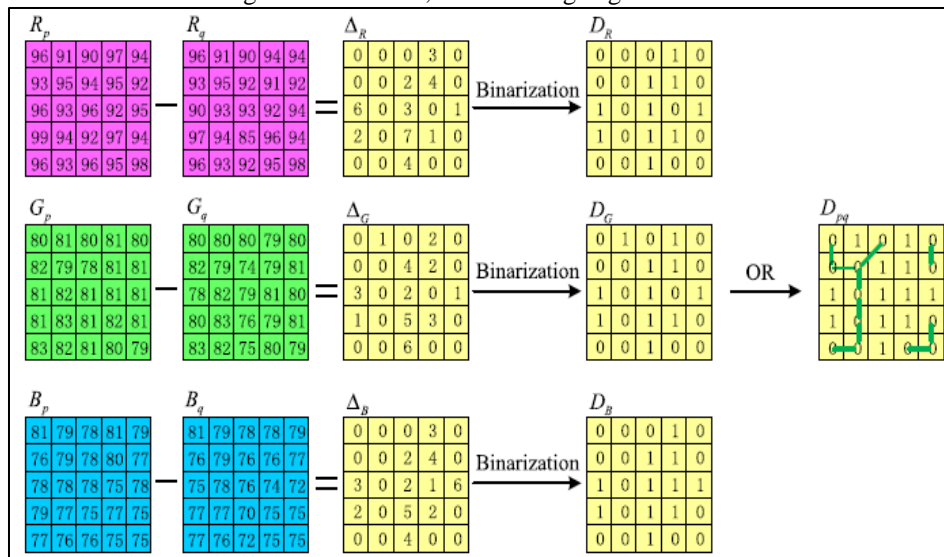


Fig. 3: Matching Degree Computation

C. Greatest Zero-Connectivity Component Labeling

After suspicious block searching, the pixels within those blocks need to be labeled to generate a group of suspicious regions. The existing methods [5,4] adopt a whole block labeling (WBL) strategy, where all pixels in suspicious blocks are marked as being tampered. However, previous shows that filling operation is only for unknown portion of the target block during the inpainting process. This means that a suspicious block contains both tampered pixels and original pixels. This phenomenon is most

pronounced on the edge of the object removal region. Hence, WBL would mistake original pixels as being tampered, thereby cause high false positive rate and get jagged edges, which will definitely decrease the precision of forgery location. In order to remove the false alarm caused by WBL and get a more accurate location of suspicious region, this work proposes a new method based on GZCL. First, we find the greatest zero-connectivity component in the difference array obtained by suspicious block searching.

D. Fragment Splicing Detection

A rough location of tampered region can be obtained using the three steps mentioned above. Nonetheless, there exists some reference regions that are falsely identified as being tampered due to the fact that matching relationships between tampered regions and its reference ones are founded during suspicious block search. Therefore, reference regions have become another cause of false alarm. In order to distinguish tampered regions from the reference ones, this work presents a method based on FSD to remove false positives caused by reference regions, producing the final location of forgery. By analyzing exemplar-based inpainting and copy move forgery, it is not difficult to find that tampered regions are replicated by multiple reference regions, in other words, more than one reference regions are matched to a tampered region. More importantly, the reference regions are distributed as scattered "fragments" and usually matched to only one tampered region. Based on this characteristic, we differentiate tampered regions from its reference ones according to the number of regions matched with a target region.

First of all, suspicious region $Region_i$ is assigned with a unique code. Secondly, the code of $Region_i$ is replaced with the code of reference region $Region_j$. The replacement of region code will help count the number of regions matched with a target region. After that a histogram was created, analysing the histogram values, the highest tampered region will considered and finally, producing the final position of tampered region.

IV. EXPERIMENTAL RESULTS

The method was implemented in a MATLAB 2014 prototype and tested with different exemplar inpainted and copy move forged images collected from various sources. The image database was built from some test images collected from various sources. In addition, the contents of those images includes landscape, buildings, human, animals, and so on. Data collection was performed with the forged images which is created by exemplar image inpainting and copy move forgery.

The proposed method is evaluated through observation and analysis to determine whether it can detect the multiple forged images accurately and efficiently. For the experimental analysis, datasets are created. The inpainted images are generated using the exemplar algorithm due to the relative better visual quality than existing methods when a large object is removed. The copy move images are created by the copying and pasting the images.

At the beginning of the process, the input image is processed through central pixel mapping. The technique CPM is a fast searching algorithm, which transforms the color information of central pixel in target block into a hash value and search for the best-match block in those blocks with similar hash values. After CPM, suspicious block searching is done. According to the filling mechanism of exemplar-based inpainting and copy move, tampered pixels will share the same values with the corresponding ones in the best-match block. i.e., the center position of the target block is the same as that of its reference block. So, the search for suspicious blocks can be achieved by only checking the blocks with same central pixels, which verifies that CPM is a proper hash function. After that, GZCL is used to mark the tampered pixels in suspicious blocks. And a statistical approach is also made to compare the tampered region with other areas. Finally, FSD and is employed to filter out the reference regions and get the final locations of tampered regions.

In addition, the proposed method is assessed through both quantitative and qualitative approach. The assessment of the results for the qualitative tests was done mainly by visual analysis. The quantitative evaluation was performed by repeating the experiments multiple times for different block size placed randomly throughout the images. These sizes were chosen as they shows variation in the output. For sake of quantitative comparison, PSNR (Peak Signal-to- Noise Ratio) between two images from Fig. 5 were computed (see Table 1). PSNR is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. PSNR values are represented in decibels (dB).



Fig. 4: (a) Original image (b) Inpainted image (c) Image inpainting detected

We have experimented with other three algorithms on some images comparing with PSNR. In this analysis, the proposed method outperforms reasonably compared to the other three methods. In the lady occluded image, the proposed method reports highest PSNR value (45.72) compared to the other three algorithms. In the man occluded image, the J. Wu et al's algorithm exhibits highest PSNR value (34.89). In the bird occluded image, the proposed method again produce highest PSNR value (36.76).



Fig. 5: Test images. (a) Lady occluded image (b) Man occluded image (c) Bird occluded image.

Table – 1
Quantitative evaluation using PSNR (in dB) for all comparative images from Fig.5

Method	PSNR value (dB)		
	Lady Occluded Image	Man Occluded Image	Bird Occluded Image
Criminisi et al's	33.12	33.16	33.22
J. Wu et al's	34.29	34.89	34.53
Z. Lu et al's	34.65	34.79	34.62
Proposed method	45.72	16.78	36.76

The comparison of time required for the proposed system corresponding to the size of the image is shown in Table 2. As shown in Table 2 we utilized four images such as house image, sealand image, lamp image and parachute image as shown in Fig. 6.



Fig. 6: Test images. (a) House image (b) Sea land image (c) Lamp image (d) Parachute image

Table – 2
Comparison of time required for the proposed method corresponding to the size of the image

Image	Size of image	Time taken (seconds)
House image	163 × 169	24
Sea land	249 × 187	158
Lamp image	195 × 131	37
Parachute image	236 × 172	36

V. CONCLUSION

Even though the exemplar-based inpainting method leaves behind no visual artifacts, but it introduces unusual similarity between image block pairs. Similarly, the copy-move forgery also introduces a co-relation between the source and the target region as they are identical block pairs. A well-organized technique is presented for detecting tampered regions from these forgeries which is based on suspicious block searching incorporates with cental pixel mapping, GZCL and FSD. The use of CPM technique makes the searching process faster hence improving the overall efficiency of the algorithm. The uniqueness of the algorithm lies in the fact that it is one of the Digital Forensic Algorithm which detects two different kinds of forgeries. We have also demonstrated the efficiency of the proposed algorithm on numerous images. The experimental results suggests that the proposed method successfully reduces the running time compared to the other existing methods. Both the qualitative and quantitative results proves that the present method outperforms reasonably compared to the existing methods.

REFERENCES

- [1] J.A. Redi, W. Taktak, J.L. Dugelay, "Digital image forensics: a booklet for beginners", *Multimedia Tools Appl.* 51 (1) (2011) 133162.
- [2] O.M. Al-Qershi, B.E. Khoo, "Passive detection of copy-move forgery in digital images: state-of-the-art", *Forensic Sci. Int.* 231 (1) (2013) 284295.
- [3] S. Bayram, H.T. Sencar, N. Memon, "An efficient and robust method for detecting copy-move forgery", in: *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 10531056.
- [4] C. Chang, J. Cloud Yu, Chih-Chuan Chang, "A forgery detection algorithm for exemplar-based inpainting images using multi-region relation" - *Image and Vision Computing* 31 (2013),pg. 5771. Elsevier Ltd.
- [5] Z. Liang, G. Yang, X. Ding and L. Li, "An efficient forgery detection algorithm for object removal by exemplar-based image inpainting", *J. Vis. Commun. Image R.* 30, 2015, 7585.
- [6] C. Barnes, E. Shechtman, D.B. Goldman, et al., "The generalized patch match correspondence algorithm", in: *Proc. of European Conference on Computer Vision*, 2010, pp. 2943.
- [7] A. Criminisi, P. Prez, K. Toyama, "Region filling and object removal by exemplar based image inpainting", *IEEE Trans. Image Process.* 13 (9) (2004) 12001212.
- [8] Q. Wu, S.J. Sun, W. Zhu, et al., "Detection of digital doctoring in exemplar-based inpainted images", in: *Proc. of IEEE International Conference on Machine Learning and Cybernetics*, vol. 3, 2008, pp. 12221226.
- [9] J.Wu and Q. Ruan, "Object removal by cross isophotes exemplar based image inpainting," in *Proc. Int. Conf. Pattern Recognition*, 2006, pp.810–813.
- [10] K.S. Bacchuwar, K.R. Ramakrishnan, "A jump patch-block match algorithm for multiple forgery detection", in: *Proc. of IEEE International Multi-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s)*, 2013, pp. 723728.