# A new method for Image Super-Resolution

Gérard LABLANCHE[1,1] and Sébastien LABLANCHE[2]

[1]Université Bordeaux 1, Bordeaux, France
[2]Université Paul Sabatier Toulouse III, Toulouse, France

[1]Corresponding author: gerard.lablanche@orange.fr/(+33) 0647665706

**Abstract**

The aim of this paper is to demonstrate that it is possible to reconstruct coherent human faces from very degraded pixelated images with a very fast algorithm, more faster than compressed sensing (CS) algorithm, easier to compute and without deep learning, so without important information technology resources, i.e. a large database of thousands training images (see https://arxiv.org/pdf/2003.13063.pdf). This technological breakthrough has been patented in 2018 with the demand of french patent FR 1855485 (https://patents.google.com/patent/FR3082980A1). The Face Super-Resolution (FSR) has many interests, in particular in a remote surveillance context which already exists in China but which can be a reality in USA and European countries. Today, deep learning methods and artificial intelligence (AI) appears in this context but these methods are difficult to put in their systems because of the need of important data. The demand of chinese patent CN 107563965 and the scientist publication "Pixel Recursive Super Resolution", R. Dahl, M. Norouzi, J. Shlens propose such methods (see https://arxiv.org/pdf/1702.00783.pdf). In this context, this new method could help governments, institutions and enterprises to accelerate the generalisation of automatic facial identification and to earn time for reconstruction process in industrial steps such as terahertz imaging, medical imaging or spatial imaging.

# Introduction

In February 2017, three researchers described before from Google Brain published their results on Pixel Recursive Super Resolution: (https://arxiv.org/pdf/1702.00783.pdf). Their technology consists in approaching the final picture by combining an algorithm and database pictures of Google to obtain from an initial 8x8 definition picture, a 32x32 definition picture which is very similar as the actual picture. The problem of this technique is mainly that the algorithm loses efficiency if the original image isn't present in the database of training images. The obtaining of high-resolution digital images can be made at the time of the shooting, but it is often synonymic of important costs because of the necessary material to avoid such costs, it is known how to use methods of super-resolution reconstruction, consisting from one or several low resolution images to obtain a high-resolution image. The american patent US 9 208 537 describes such an algorithm. A zone of one low-resolution image is isolated and categorized according to the information contained in pixels forming the borders of the zone. The category of it zone determines the type of interpolation used to add pixels in aforementioned zone, to increase the neatness of the images. But the problem is the high complexity of this class of algorithm. A French startup is said to have developed an innovation based on a powerful algorithm which could deliver similar results as Google, BUT, without using any database. It would represent a very important innovation for any potential end user which would want to operate independently of Google database, and faster. In this paper, we present the new technical of interpolation for image super-resolution based on conditional interpolations in three directions called successively to achieve a blurred high resolution image with a lot of more details to help deconvolution algorithm (the directional interpolation monitored by the orientation of the motion filter) to create truth details at the end (LABLANCHE process). The strange thing in this discovery is that the only one filter which finds human faces is the motion filter whereas the gaussian filter fails, so the scientists and scholars will can extrapolate different interpretations of this technique in other scientific domains.

# Model

The optimization model is built with six interdependant parameters: p2, p3, p4, $\gamma$, L and $\theta$. The first three p2, p3 and p4 are parameters of LEVEL 1 conditional interpolations described in the demand of french patent FR 1855485 (parameters chosen between 0 and 255).
$\gamma$ is the magnification factor of the image to determine for the test image before the deconvolution and (L, $\theta$) is the couple (distance, angle) of the motion filter.
The optimal image is

$$\hat{x} = \underset{p2,p3,p4,\gamma,L,\theta}{\arg\min} \|Ax - b\| + \lambda \|x\|_1 \tag{1}$$

where A is the mixture (deconvolution, conditional interpolation, magnification factor) reconstruction matrix, b the data of pixelated image, and $\|x\|_1$ the $l_1$ minimization in an appropriate sparse basis (i.e. 9/7 wavelet) and lambda is the compromise between the two terms.
B1 (left), B2 (right), B3 (left and bottom) and B4 (right and bottom) are the four neighbor blocs of the B bloc of 16 pixels on the image (see Figure 1).
B1 keeps the color of B bloc (always),
B2 takes the average color $\frac{B+C}{2}$ if $|B - C| <$ p2 and keeps the color of B bloc otherwise,
B3 takes the average color $\frac{B+D}{2}$ if $|B - D| <$ p3 and keeps the color of B bloc otherwise,
B4 takes the average color $\frac{B+E}{2}$ if $|B - E| <$ p4 and keeps the color of B bloc otherwise.
p2 is the parameter which controls the mean between the B bloc and the C bloc.
p3 is the parameter which controls the mean between the B bloc and the D bloc.
p4 is the parameter which controls the mean between the B bloc and the E bloc.
The setting of p2, p3 and p4 parameters is very important. We realize the interpolation between the two blocs only if they have enough closed colors. These parameters are integers between 0 and 255. If the value is equal to 0, there is no interpolation and if the value is 255, the interpolation is always realized. In other cases, the interpolation is realized if the two adjacent blocs have intensities with a smaller difference than the constraint. The constraints p2, p3 and p4 define a metric between two adjacent blocs in three directions. The goal of this step is to separate different interest areas of the image and to generate contrast enhancement.

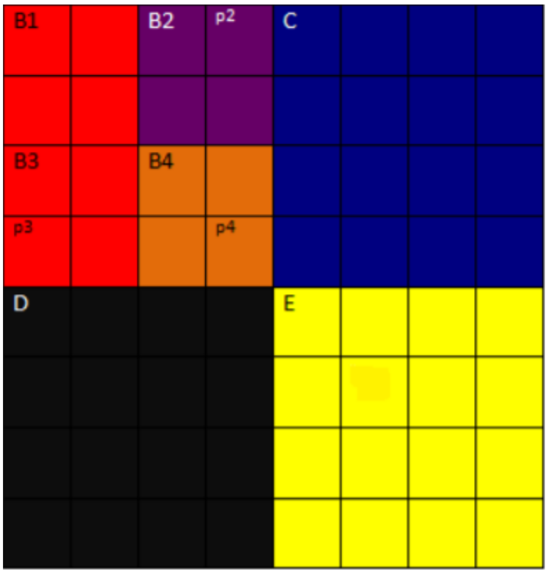


Figure 1: B1,B2,B3 and B4 blocs (LEVEL 1)

# Results and discussion

We use 10 test images of 32x32 pixels for our algorithm: Google Brain (x3), marie bonneau (x2), Ellie Goulding, Ariana Grande, Man, Shailene Woodley and Ronald Coifman (eye).

Figure 2: Google Brain

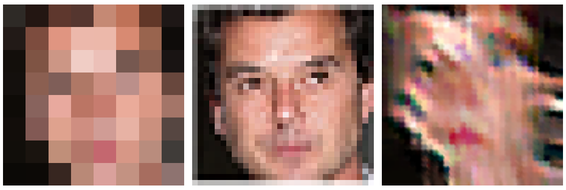

Figure 3: Google Brain

Figure 4: Google Brain

Figure 5: marie bonneau

Figure 6: marie bonneau

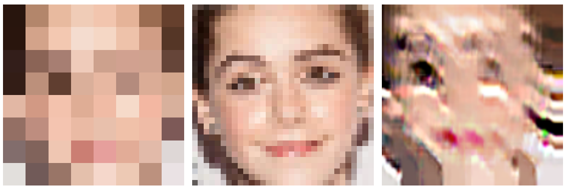

Figure 7: Ariana Grande

Figure 8: Ellie Goulding

Figure 9: Man

Figure 10: Shailene Woodley

Figure 11: Ronald Coifman (eye)

## Discussion

1. First advantage: Our algorithm doesn't need CUDA graphics card to run very fast because the setting of parameters is already known into the program and the complexity of conditional interpolations is very low.

   ```
   Dichotomy61_IMAGE(nbl, nbc, v, step, q, p2, p3, p4); \\p2, p3 and p4 are known
   \\only one call of the routine
   ```

2. Second advantage: The learning of parameters is already known for human faces and the natural aspect of the reconstruction is achieved without training images database.

| | |
|---|---|
| Google Brain | L=15,$\theta$=90° |
| Google Brain | L=15,$\theta$=70° |
| Google Brain | L=17,$\theta$=115° |
| marie | L=13,$\theta$=105° |
| marie | L=12,$\theta$=160° |
| Ariana Grande | L=17,$\theta$=95° |
| Ellie Goulding | L=11,$\theta$=125° |
| Man | L=20,$\theta$=140° |
| Shailene Woodley | L=13,$\theta$=120° |
| Ronald Coifman (eye) | L=10,$\theta$=170° |

Table 1: deconvolution parameters (right images)

3. Third advantage: This model is the most robust in the state of the art because with only six parameters (p2, p3, p4, $\gamma$, L, $\theta$) we achieve in average between 90% and 100% of AI details results.

4. Negative part: Our algorithm gives very good results with 4x4 pixels blocs but catastrophic results with 2x2 blocs and 3x3 blocs, so the camera system need to generate pixelated images with 4 by 4 pixels blocs.

5. Negative part: The color of eyes and the size of cheeks are the two details impossible to recover exactly and we note the presence of occlusions because of the strong deconvolution to apply.

# Conclusion

In this paper we introduce a new class of algorithms based on new interpretations of deconvolution. The motion deconvolution is used to reconstruct details of the image instead of deblur forensic images. We note that the complexity is low and this process is easy to compute in comparison with deep learning, compressed sensing, mixture of both or categorization of zones. Object motion blur is caused by the relative motion between an object in the scene and the camera system during the exposure time. This type of blur generally occurs in capturing a fast-moving object or when a long exposure time is needed. If the motion is very fast relative to the exposure period, we may approximate the resultant blur effect as a linear motion blur, which is represented as a 1D local averaging of neighboring pixels and given by

$$h(i,j,L,\theta) = \begin{cases} \frac{1}{L} & \text{if} \quad \sqrt{i^2+j^2} \leq \frac{L}{2} \quad \text{and} \quad \frac{i}{j} = -tan(\theta), \\ 0 & \text{otherwise,} \end{cases}$$

where (i,j) is the coordinate originating from the center of h, L the moving distance and $\theta$ the moving direction (or the angle). In practice, however, real motions are extremely complex and cannot be approximated by such a simple parametric model. An appropriate way to handle this issue is to use the non-parametric model, i.e. no explicit shape constraint is imposed on the blur kernel, and the only assumption is that the kernel needs to follow the motion path. A more serious issue is in an image where only the region of moving objects is disturbed by the blur kernel, while other regions remain clear. The future research on this topic would be to analyse the behaviour of the motion blur in the different areas of the image.

# Acknowledgments

We lead this work with the background of Université Paul Sabatier Toulouse III (Toulouse, France). Sébastien LABLANCHE is graduated with MASTER 1 in signal processing at Université Bordeaux 1 in 2010 and MASTER 2 in statistics at Université Paul Sabatier Toulouse III in 2012.
We thanks to Jean-Marc AZAIS (IMT) and Charles DOSSAL (IMB).

## References


[1] Yonina C.Eldar and Gitta Kutyniok, *Compressed Sensing Theory and Applications*, Cambridge.

[2] Stephane Mallat, *Une exploration des signaux en ondelettes*, les éditions de l'école polytechnique.

[3] Charles Dossal, *Estimation de fonctions geometriques et deconvolution*, These pour l'obtention du titre de Docteur de L'Ecole Polytechnique, le 5 Decembre 2005.

[4] Richard E.Blahut, Willard Miller, Jr, Calvin H.Wilcox, *Radar and Sonar Part I*, Springer-Verlag.

[5] Sung Cheol Park, Min Kyn Park and Moon Gi Kang, *Super-Resolution Image Reconstruction: A Technical Overview*, IEEE SIGNAL PROCESSING MAGAZINE, 1053-5888/03, May 2003.

[6] Pradeep Sen and Soheil Darabi, *Compressive Image Super-resolution*, Advanced Graphics Lab, Department of Electrical and Computer Engineering, University of New Mexico, USA, New Mexico, 2009 IEEE,1235-1242.

[7] Jianchao Yang, John Wright, Thomas S.Huang, Yi Ma, *Image Super-Resolution via Sparse Representation*, IEEE Transactions on image processing, vol.19, pp. 2861-2873, November 2010.

[8] Michael Elad and Michal Aharon, *Image Denoising via Sparse and Redundant Representations over Learned Dictionaries*, IEEE Transactions on image processing, vol.15, pp. 3736-3745, December 2006.

[9] Michal Irani, Shmuel Releg, *Image Sequence Enhancement using multiple motions analysis*, Institute of computer Science, The Hebrew University of Jerusalem, 0-8186-2855-3/92, 1992 IEEE.

[10] John Wright, Allen Y. Yang, Arvind Ganesh, S.Shankar Sastry, Yi Ma, *Robust Face Recognition via Sparse Representation*, February 2009 IEEE Transactions on pattern Analysis and Machine Intelligence, vol.31, pp. 210-227.

[11] Hong Chang, Dit-Yan Yeung, Yiming Xiong, *Super-Resolution Through Neighbor Embedding*, Hong-Kong University of Science and Technology, IEEE 2004, Computer Society Conference on Computer Vision and Pattern Recognition, vol.1, pp. 275-282.