Enhanced Resolution Methods for Improving Image Analysis and Pattern Recognition in Scanning Probe Microscopy

Mario D'Acunto[1], Gabriele Pieri[2], Marco Righi[2] and Ovidio Salvetti[2]

[1]Institute of Structure of Matter, National Research Council, ISM-CNR, via Fosso del Cavaliere 100, I-00133, Rome, Italy
[2]Institute of Information Science and Technology, National Research Council, ISTI-CNR, via Moruzzi 1, I-56124, Pisa, Italy

Abstract. Image acquisition systems integrated with laboratory automation produces multi-dimensional datasets. An effective computational approach to objectively analyzing image datasets is pattern recognition (PR), i.e. a machine-learning approach where the machine finds relevant patterns that distinguish groups of objects after being trained on examples (supervised machine learning). In contrast, the other approach to machine learning and artificial intelligence is unsupervised learning, where the intelligent process finds relevant patterns without relying on prior training examples, usually by using a set of pre-defined rules. In this paper we apply a method derived by usual PR techniques for the recognition of artifacts and noise on images recorded with Atomic Force Microscopy (AFM). The advantage of automatic artifacts recognition could be the implementation of machine learning languages for AFM investigations.

1 Introduction

It's important for machine image understanding to have high resolution images and to recognize the semantic of the image (in other word what are the represented object and which is their sense). In our work, we study super resolution (SR) algorithms in order to have a high information density from our data and we apply PR algorithms on high resolution images in order recognize the features of the analyzed images [1,11,13].

Within the field of image analysis applied to screening device, this paper will focus on the direct correlation between SR methods and PR methods. In particular, SR algorithms will be used to recognize patterns of device recorded images and provide an accurate feedback for checking real time device operability, i.e. using machine learning algorithms.

SR algorithms generate a denoised hyperesoluted image (or a set of images) from low resolution ones. The knowledge of the class of images to analyze helps during the computation of the high resolution image. The higher information contained in the generated image provide a better sample in order that can be easily use by pattern recognition algorithms. The results provided from the PR algorithm supply a data

D'Acunto M., Pieri G., Righi M. and Salvetti O..
Enhanced Resolution Methods for Improving Image Analysis and Pattern Recognition in Scanning Probe Microscopy.
DOI: 10.5220/0004392400220028
In *Proceedings of the 4th International Workshop on Image Mining. Theory and Applications* (IMTA-4-2013), pages 22-28
ISBN: 978-989-8565-50-1

input for the machine learning algorithms that gives the possibility to change the device regulation in order to obtain better images.

This operative procedure can be applied to a big set of devices used for automated data acquisition. In fact the environments with a high grade of automation produces a huge image dataset that can be hardly hand checked so it is necessary provide an automated control system that can provide an intelligent feedback to the devices in order to maintain the devices to the highest efficiency.

A particular and innovative application field is the Scanning Probe Microscopy (SPM). In fact the advent of SPM family instruments since the 80 decade of the last century opened the possibility to observe and manipulate matter at atomic scale making possible to improve the knowledge and technology on nanoscale (commonly claimed Nanotechnology). Nevertheless, today the application of SPM techniques is limited by the fact that the experimental scanning best conditions can be found only manually.

After a theoretical study of the pipeline composed by SR algorithms, PR algorithms and device control by the mean of artificial intelligent algorithms we focused our work in a possible application on a SPM device.

2 An Overview on PR Features

Observing an image a human can notice some particular pattern or characteristics that are unique for a certain type of material. This inference process is useful in order to observe phenomenon and so on. For example, in computer vision a computer can only analyze a set of matrix (one or more matrix) for each image in which the colors are coded using a particular color code. The data extraction performed by a computer and its interpretation is a task that permits to a machine to recognize patterns, regularity and provide an interpretation.
The main approach to patter recognition can be classified as follow:

- statistical learning
- classification

The statistical learning is extremely important as show in numerous examples [11] such as predict the price of a stock six mounts from now, estimate if the received e-mail is or is not spam, recognize handwritten characters and digit and understand if an image contains archaeological handmade objects [12]. A first kind of classification can divide learning problems into two sets: supervised and unsupervised. In supervised learning an algorithm provide a predicable output based on a set of input measures, in unsupervised learning the algorithm objective is “understand” the relation between a set of input (i.e. analyzing recurrent pattern).
The classification works on predictors $S(x)$ which takes value in a discrete set S. Usually the input space is divided into some labeled regions according to the *S-classification*. The boundaries between the regions can be of roughed or smoothed. For each x_i given as input, the classifier provide a g_i as output, where $g_i \in S$. There are some methods to determine g_i: prototype, *K*-means clustering, learning vector quantization, *K*-nearest neighbors, neural networks, kernel methods and support vector machines [13].

3 SR Methods for Improving PR

The resolution of an image is determined by many factor depending on the acquisition system. The equation 1 describes the imaging model we use.

$$g(x, y) = h(x, y) * f(x, y) + n(x, y). \quad (1)$$

In details, $h(x, y)$ is the point spread function (PSF), $f(x, y)$ is the ideal image, $g(x, y)$ is the original image and $n(x, y)$ is the noise.
The problem of the resolution in AFM images depends on tip control and feedback. The main factor that determine the resolution of an image is the number of pixels that describes an area in the real image [1-5]. The increasing of pixel size is not only a pleasure but can reveals important particulars. The SR techniques that can be classified into two classes: single-frame image restoration algorithms or multi-frame image restoration algorithms.
The classic algorithms get a single image in input and produce a single output image. The introduction of digital video, i.e. by the means of surveillance camera, led to the analysis of multi-frame images. Even if in a video each frame represent different images, consequentially frames are quite similar so that it is possible using them in order to process the data.
The video analysis conduct to study techniques of motion estimation. Following this research field, it was recognized the potential of image restoration in order to increase the spatial resolution using similar images. The application of motion compensations and image restoration algorithms in order to produce high-quality and high-resolution still images conduct to the so called super-resolution reconstruction (SRR).
The SRR algorithms transform low resolution images into a high resolution image. In order to produce the high resolution image it is necessary to remove the effects of possible blurring and noise from the low resolution images. In other words, the SRR algorithm computes low resolution images by blur, noise and aliasing [1-3],[6-7].
The SRR algorithms are applied to a large number of problems such as satellite imaging, astronomical imaging, video enhancement [8-9] and restoration, microscopy [10] and other.

The algorithm we study comes from an idea suggested by Zou [14]. This SR algorithm uses a training set in order to get a high resolution image of a face. For example, it is able to transform a low 16x12 pixel image into a high 64x48 pixel image. During the training, the algorithm get in input one set of low resolution images and one set of high resolution images (for each low resolution image there is a high resolution image). By the computation of this training set the algorithm generate a set of rules in order to transform new low resolution images into high resolution images.

We test the algorithm using Yale and Feret archives performing two sets of independent tests. In figure 1 we show the result of our test, the picture A shows the source image, the picture B shows the output of our algorithm that we can compare with the picture C that is obtained with a bi-cubic interpolation, finally, the picture D shows the original high resolution images.

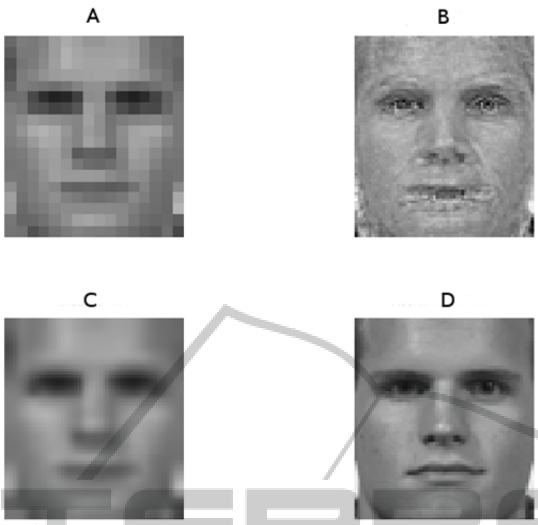

Fig. 1. A face is used during the algorithm test. The picture A shows the source image, the picture B shows the output of our algorithm that we can compare with the picture C that is obtained with a bi-cubic interpolation, finally, the picture D shows the original high resolution images.

4 AFM Imaging Improved by PR Methods Combined with SR: Numerical Results

Among the SPM family, AFM operating mechanism based on sensing the specimen through the force between its surface and a sharp probe. A cantilever oscillates and touches the biological sample only intermittently at the end of its downward movement, which reduces the contact time and minimizes friction and destructive forces. This is why AFM produces high-resolution topographic and force measurements in aqueous and physiologically relevant environments without the need to stain or pre-treat the specimens.

The most important advantage of applying AFM in biological research related to the fact that AFM is essentially a single-molecular technique, providing insight into the geometry, elasticity and dynamic behavior at the level of single molecular or single cell. As many biological processes, such as protein amyloid self-assembly, involve multiple pathways and are characterized by inherent heterogeneity of species, the application of single molecule studies is of critical significance.

Preliminary results following the algorithm described in the preview section. During our first experiments we take in input the picture A of the figure 2 (picture B shows the 3D aspect of the surface). As a result of data process, we have a well characterized profile of the surface as showed in picture C (picture D shows the 3D aspect of the surface).

The great advantage of AFM is that the screening procedure over large number of potential partners can be carried out in their natural environment without their pretreatment or fixation. This non-invasive procedure can be applied for identification

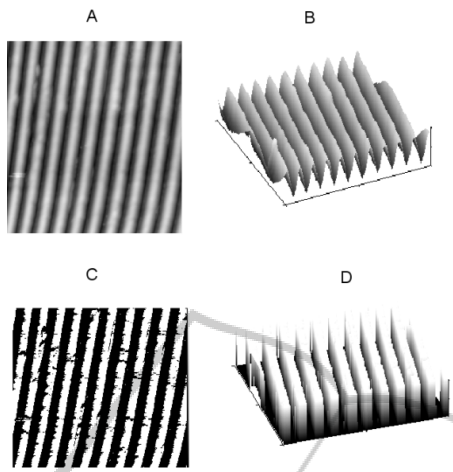

Fig. 2. From A to B: The input image and its 3D rendering. From C to D: the output after our data processing.

of the promising lead compounds among the large library of biological active species, which would display the largest attractive forces towards their target molecules.

Our approach for improving single image PR on biological samples is based on the following steps, first a standard PR method is applied to an image in order to define the image features. The second step regards the increasing of pixel density on the image using SRR approach and finally, the third step is devoted to pattern matching between the first image and the enhanced image.

An example of the application of our approach to biological sample is shown in figure 3. The image of a fibroblast cell is processed following the above described sequence. The pattern to be recognized are inherent the specific intra-cell organs included subsurface actins and filaments.

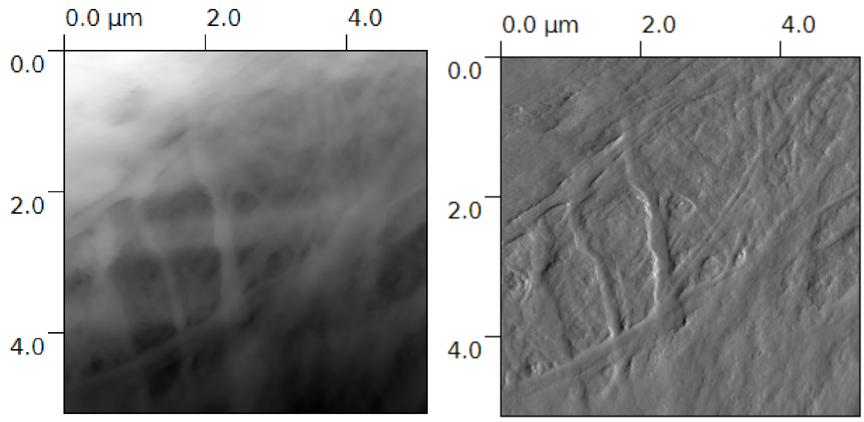

Fig. 3. On the left we have a low resolution image 5μm×5μm of a fibroblast cell as recorder by an AFM and processed with commercial software (Park Scientific Instruments) and free available software (Gwyddion). On the right, we have the correspondent SRR image. Now, on the right image it is possible to estimate the different cytoskeleton components, as actin and filaments (dimensions approximately 100nm).

Figure 3 summarizes the effective advantage by using our algorithms. On the left, we have a low resolution image 5µm×5µm of a fibroblast cell as recorder by an AFM and processed with commercial software (Park Scientific Instruments) and free available software (Gwyddion). On the right, we have the correspondent SRR image. From the initial image (on the left) it is possible to have an idea of the various cytoskeleton cell organs, but the low quality image makes difficult to estimate the plot of such organs and their real dimensions. On the contrary, improving the pixel density in a reasonable way using SR methods, it is possible estimate the cytoskeleton plot and to identify the organs with their real dimension, approximately 100nm (1nm=1nanometer=10^{-9}meter).

5 Conclusions and Future Perspectives

In this paper, we focused the attention on an effective computational approach to increase the resolution of Scanning probe Microscopy image for improving the pattern recognition. The results obtained can be considered as a first step of a more general framework for applying machine learning and artificial intelligence to nanoscale imaging, where the intelligent process finds relevant patterns without relying on prior training examples, usually by using a set of pre-defined rules. In details, we apply a method derived by usual pattern recognition techniques for the recognition of artifacts and noise on images recorded with Atomic Force Microscopy. First immediate advantage of such automatic artifacts recognition could be the implementation of machine learning languages for AFM investigations.

Acknowledgements

The authors wish to thank the NanoICT Project for useful support.

References

1. Image Process. 6: 774-8. Kang, M. G.; and Chaudhuri, S. 2003. Super-resolution image reconstruction. IEEE Signal Process. Mag. 20: 19-20.
2. Ng, M. K.; and Bose, N. K. 2003. Mathematical analysis of super-resolution methodology. IEEE Signal Processing Magazine 20(3):62-74.
3. Park, S. C.; Park, M. K.; and Kang, M. G. 2003. Super-resolution image reconstruction: a technical overview. IEEE Signal Processing Magazine 20(3): 21-36.
4. Rajan, D.; Chaudhuri, S.; and Joshi, M. V. 2003. Multi-objective super resolution concepts and examples. IEEE Signal Processing Magazine 20(3): 49-61.
5. Rajan, D.; and Chaudhuri, S. 2003. Simultaneous estimation of super-resolution scene and depth map from low resolution defocuses observations. IEEE Trans. On Pattern Analysis and Machine Intelligence 25(9): 1102-17.
6. Processing 5(6): 996-1011. Segall, C. A.; Molina, R.; and Katsaggelos, A. K. 2003. High-resolution images from low-resolution compressed video. IEEE Signal Processing Magazine 20(3): 37-48.

7. Callicó, G. M.; Núñez, A.; Llopis, R. P.; and Sethuraman, R. 2003. Low-cost and real-time super-resolution over a video encoder thIP. Proc. 4 IEEE Int. Symp. on Quality Electronic Design (ISQED'03), San Jose, CA, USA, 24-26 March 2003, pp. 79-84. IEEE Computer Society, Washington, DC, USA.
8. Jiang, Z.; Wong, Tien-Tsin; and Bao, H. 2003. Practical super-resolution from dynamic video sequences. Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'03), Madison, WI, USA, 16-22 June 2003, vol. 2, pp. 549-54. IEEE Computer Society, Washington, DC, USA.
9. Zibetti, M. V. W.; and Mayer, J. 2005. Simultaneous super-resolution for video sequences. Proc. IEEE Int. Conf. on Image Processing (ICIP'05), Genoa, Italy, 11-14 September 2005, vol. 1, pp. 877-80. IEEE Signal Processing Society, Piscataway, NJ,USA.
10. Karthik Kumar1, Huigao Duan1, Ravi S. Hegde, Samuel C. W. Koh1, Jennifer N. Wei and Joel K. W. Yang; Printing colour at the optical diffraction limit; august 2012, pp. 557-561, Nature Nanotecnology.
11. T. Hastie, R. Tibshirani, and J. Friedman. The elements of Statistical Learning - Data Mining, Inference and Prediction. Springer, 2001.
12. Benedetto Allotta, S. Bargagliotti, L. Botarelli, Andrea Caiti, Vincenzo Calabro, G. Casa, M. Cocco, Sara Colantonio, Carlo Colombo, S. Costa, M. Fanfani, L. Franchi, P. Gambogi, L. Gualdesi, D. La Monica, Massimo Magrini, Massimo Martinelli, Davide Moroni, Andrea Munafò, Gordon J. Pace, C. Papa, M. A. Pascali, Gabriele Pieri, M. Reggiannini, M. Righi, Ovidio Salvetti, Marco Tampucci: Thesaurus Project: Design of New Autonomous Underwater Vehicles for Documentation and Protection of Underwater Archaeological Sites. EuroMed 2012: 486-493.
13. J. Shawe-Taylor and N. Cristianini. Kernel methods for pattern analysis. Cambridge University Press, Cambridge, 2004.
14. W. W. W. Zou and P. C. Yuen; "Very Low Resolution Face Recognition Problem"; Transactions on Image Processing, vol. 12, N. 1, January 2012