High Resolution Images from Low Resolution Video Sequences

Federico Cristina
fcristina@lidi.info.unlp.edu.ar - *Ayudante Diplomado UNLP*

Sebastián Dapoto
sdapoto@lidi.info.unlp.edu.ar - *Becario III-LIDI*

Claudia Russo
crusso@lidi.info.unlp.edu.ar - *Profesor Adjunto Dedicación Exclusiva*

Oscar Bria
onb@lidi.info.unlp.edu.ar - *Profesor Adjunto Dedicación Exclusiva*

*III-LIDI – Instituto de Investigación en Informática LIDI.
Facultad de Informática. Universidad Nacional de La Plata.*

Abstract

In some cases, low resolution of those images composing a video film hinders the proper visual interpretation of its data. A typical example of this is video obtained from security cameras.

There thus exists the need to count with some method allowing the processing of such information in order to obtain a better quality and a higher level of detail of those images.

This gives rise to the possibility of making a more reliable interpretation of images, all of which eases the determination of, for example, some people face features or a car plate numbers.

Nowadays, there exist some techniques that are related to this topic (called Image Super-Resolution techniques), though in the theoretical field in principle. Besides, there is no integral solution presented as integral product for its utilization.

This paper presents the preliminary results of the Super-Resolution techniques applied to video sequences with the possibility of using quality enhancement preprocessing in each individual image.

Key Words: Super Resolution, Image Enhancement, Image Processing

Introduction

When taking an image with a digital camera, or digitizing a video sequence, the following problem arises: the information that we photograph has to be discretized and reflected in pixels so that it can be represented in a computer. We thus lose both spatial information (we take a real life image into a discrete and finite pixel grid) and information of each pixel intensity (we take the brightness and color levels to, for example, a scale of 256 intensity levels in gray scale).

Graphically, the following images show a simulated example
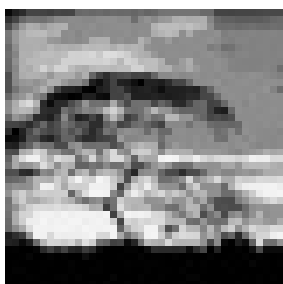

Figure 1. Image to be photographed

Figure 2. Digitalized Image

Let Figure 1 be a version to be photographed (this is only an assumption since it is obvious that this has already been digitalized). Notice how the information has been lost in both senses (space and intensity of each pixel)

Even though the example has been taken to the extreme, it is not unusual that video film images taken with security cameras are of poor quality. To this problem some other complications are added, such as noisy images, blurred images, out of focus, etc.

Enhancing an image

The simplest way of obtaining a basic enhancement in an image resolution is applying this image to some of the so-called interpolation techniques. The most popular are bicubic, bilinear, and the nearest neighbor technique (here mentioned decreasingly, taking into account the quality of the result obtained). Even though such methods present a fast solution, this is not enough in surveys, in which the certainty of the observed information in the image must be the highest.

Although there exist some other methods that allow achieving an even greater enhancement, there are yet not enough to the effects of obtaining a significant optimization.

Taking advantage of an image sequence

When we count with a video sequence where the information of a frame and the following is almost the same (i.e., the captured movement throughout the video sequence is relatively smooth), we can achieve an improvement in the problem previously described.

This is achieved by means of a technique called Image Super-Resolution [1], [2], [7], [14], which is based on taking advantage of non-redundant information of a video sequence in order to obtain as result an image of higher resolution.

Super-Resolution allows minimizing the discretization problem and the quantification error. The first problem can be summed up as the dilemma of determining in which pixel certain part of the photographed image should be placed; whereas the second problem is presented when we have to decide how intense such pixel should be, taking into account that we have a finite number of values that can be assigned to.

Figure 3 depicts the first problem. We can easily determine that the gray point must be stored in position (3,4) of the pixel matrix that make up the image. But, where should the black point be stored? (In practice, a quite used solution to this problem is to encompass more than one pixel in order to represent the point. In the example just mentioned, this would mean reflecting the black point observed in positions (1,1); (1,2); (2,1); (2,2) of the captured image with a lesser intensity than the real one. From this, the concept of Point Spread Function or PSF arises, which refers to how much the fraction of the observed target will influence the neighbor area of the pixels representing such fraction in the image once it is digitalized).
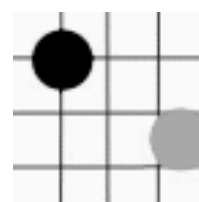

Figure 3. Two points that should be reflected in the image

Figure 4 shows the quantification problem. Assuming that a pixel (x,y) of an image deserves an intensity level of 122.6, such pixel will have to be stored with level 123, though this may not correspond to the observed value.
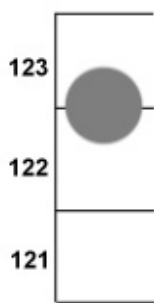
Figure 4. A pixel light intensity

Let's see now how Super-Resolution techniques can help us with the problems presented in this paper.

Basically, the proposed solution is to generate a mean of the most correct position in which a pixel represents a fraction of the real image. In consequence, if there are 3 images similar to Figure 3, where in the first the black point is stored in position (1,1), in the second in position (1,2), and in the third in position (2,1), we can determine that the optimal solution is to store the black point in position (1,1).

On the other hand, a media of each pixel can be obtained between the images that make up the video sequence. In this way, if there are 3 images, in which the pixel (x,y) in question has an intensity value of 123 in the first image, 122 in the second, and 124 in the third, we will observe that the most correct value for this will be of 123.

Though presented in a simplified way, the Super-Resolution basis is the previously described. The initial difficulty lies in the fact that the filmed object movement in the video sequence is not generally smooth, or that the camera itself is the one which generates movement in the objective. The problem in this assumption is to determine which pixels of a frame correspond to the pixels of the previous frame.

Now, here appear the motion compensation techniques [9], [10], [11], which are used together with those of Super-Resolution. Thanks to these, we can map the pixels of those images contained in a video with respect to a reference image of this.

Figure 5 shows how the motion compensation technique works: it determines that pixel (1,1) of the first image has moved towards position (3,3) in the second one. Movement vectors of each pixel between an image and the other are thus obtained.
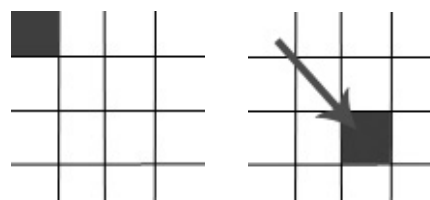

Figure 5. A pixel (x,y) movement compensation between the first and second image.

There remains another complication. Generally, video films have blurred images, with few contrast, with noise (periodic and random), and even images with part of the objective of interest blocked.

The solution to these problems is to carry out an individual preprocessing of each image before applying motion compensation and Super-Resolution. Such preprocessing should be carried out according to the characteristics of the video we are going to deal with. That is why this cannot be carried out automatically, and it implies a custom-designed task; anyway, here we will try to generate an integral solution encompassing the three previously described aspects.

In brief, the main steps in the resolution enhancement of an image from a video sequence are the following:

1. Individual preprocessing of each image according to the problem they present (blur, noise, etc).
2. Movement vector estimation between a referential image belonging to the sequence and the remaining images.
3. Application of Super-Resolution with the "clean" and "compensated" images.

Techniques studied to the present

In order to achieve an optimal result, both preprocessing and motion compensation and Super-resolution techniques are being studied concurrently. We shall see now some of the results obtained:
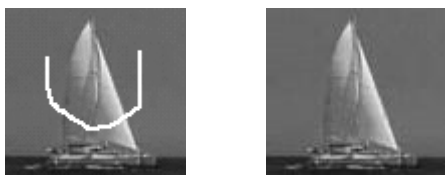
· Preprocessing

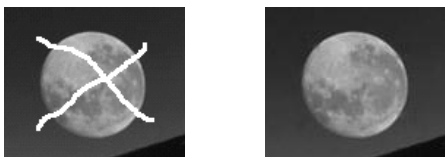Figure 6a. Inpainting

Figure 6b. Inpainting

Figure 7. Histogram Equalization

Figure 8. Averaging

Figure 9. Local Enhancement

Figure 10. Blur from Deconvolution

· Motion Compensation

Figure 11a. Block Matching Motion
Compensation (linear motion)

Figure 11b. Block Matching Motion
Compensation (rotational motion)

Figure 11c. Block Matching Motion
Compensation (deformation)

· Super Resolution by POCS

Figure 12. Bicubic interpolation vs. Super
Resolution technique by POCS
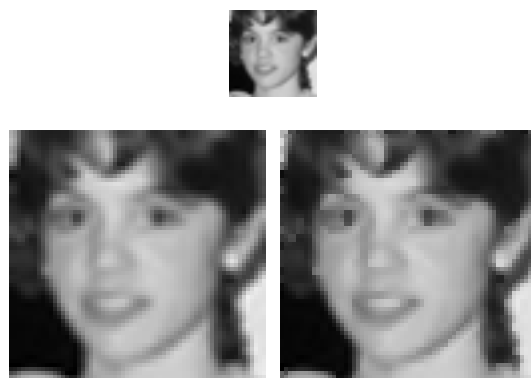
· Bayesian Interpolation

Figure 13. Bicubic interpolation vs. MAP
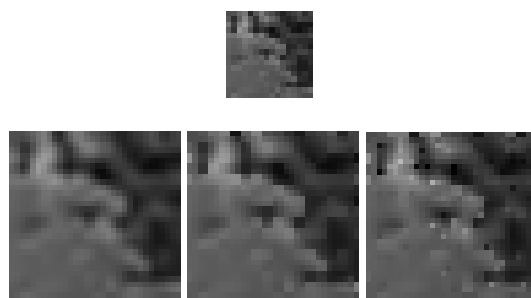Interpolation.

· Super Resolution by MAP

Figure 14. Bicubic Interpolation vs. MAP
Interpolation vs. Super Resolut. MAP

Figures 6a and 6b show the "inpainting" or "desocclusion " technique [15], [17], which regenerates the part of the image that is lacking from the neighboring information.

Figure 7 shows the histogram equalization technique [16], which helps highlighting the areas of low contrast.

Figure 8 applies the averaging technique, which allows eliminating noise when there is more than one image.

Figure 9 is based on the media and the standard deviation of the complete image and of each pixel to be studied, allowing the detection of slight changes in the image [16].

The technique used in figure 10 allows us to eliminate an image blur by means of an unsharp

mask, taking into account the Point Spread Function.

Figures 11a, 11b and 11c show the technique of Block Matching Motion Compensation in three different situations: linear movement, rotational movement, and object deformation, respectively.

Figure 12 shows the application of the technique of Super Resolution based on Projection Onto Convex Sets or POCS for a set of 6 images similar to the upper one. The bottom left image shows the result by means of the application of bicubic interpolation; while the bottom right shows the result when applying Super Resolution POCS.

Figure 13 shows the results obtained applying the technique called Maximum A Posteriori (MAP) [5], [6], [13], [18], which is based on the bayesian theory. In this case, we have used a single low resolution image in order to carry out the interpolation.
Super-Resolution by means of MAP is remarkably superior with respect to the results compared to the technique POCS.

Figure 14 shows a comparison between the bicubic interpolation, the bayesian interpolation and the initial results obtained by Super Resolution MAP, using a sequence of 9 similar low resolution images. Though almost unnoticeable, we can see the presence of "peaks" in the image obtained by Super Resolution MAP, which could not be recovered by either of the other techniques.

Proceedings were carried out in real cases applying the previously described techniques in order to obtain improvements in the quality of the photos, thus achieving the optimization desired by the user.

Research Future

In the MAP technique, an initial estimation of the enhanced image is carried out, using the Huber-Markov Random Field or HMRF

as a priori model [3], [4], [8], [12]. Later, such image is optimized in successive iterations.

The enhancement is possible by combining the a priori estimation and an equation system which relates the a priori approximation to the low resolution images that make up the video sequence.

This allows us to solve the problem in which the equation system resolution has more than one solution. This is the technique which is currently being under research. Even though it is a method that requires a great processing capacity, the difference in the results is really outstanding.

Super Resolution by MAP uses a motion compensation system called Hierarchical Subpixel Motion Estimation or HSME, since it assumes that the movement of an object in a video sequence is not necessarily at pixel level.

Hence, we have to distinguish the difference between the concept of movement at pixel level (the movement carried out by an object between two images, which may be registered in one or more pixels of distance between the first and second position) and the movement at subpixel level. In this case, the movement is very slight, which, being the image a discrete matrix, cannot reflect such change (this bounded, in turn, by the image resolution), turning this movement into a change in intensities (of gray, for instance) of the object in question between the images.

Figure 15 explains in a simplified manner the concept already mentioned. The upper left image shows a point to be digitalized and the upper right image shows such point already captured in position (1,1). If we assume that in the following frame the point has moved as indicated in the bottom left image, when reflecting such point in the pixel grid that make up the image, this should be done by storing it in positions (1,1); (1,2) and (2,2); thus obtaining an approximated representation of the new position of the point, just like the bottom right image shows. This representation will be also affected by other part of the objective captured, as can be observed in Figure 16.
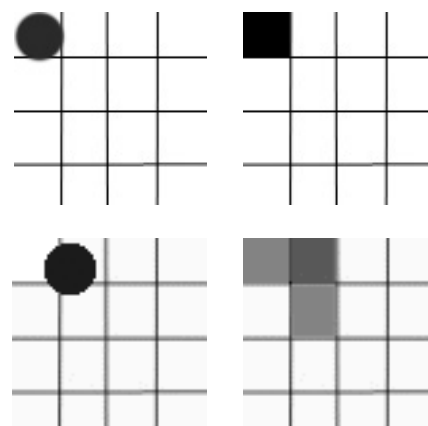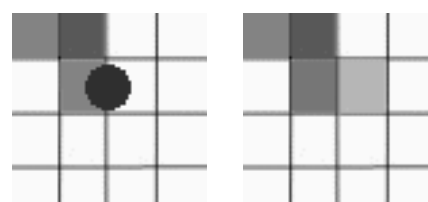

Figure 15. Subpixel Movement & PSF

Figure 16. Influence of PSF

Taking into account both the object movement between frames at pixel level and subpixel level is crucial to achieve a final optimal result.

Expected Results

The initial aim of the research is to develop an algorithm of Super Resolution by MAP applying HSME. Here we priorize the quality of the results over performance.

The following stage would encompass the integration of Super Resolution developed in the previous stage with the processing techniques of individual images that make up a video sequence.

The final stage of the research would cover the development of an integral and optimized software solution encompassing the techniques presented here in order to count with a complete product, accessible to end users.

Bibliography

[1] *Spatial Resolution Enhancement of Low - Resolution Image Sequences. A Comprehensive Review with Directions for Future Research.*
Sean Borman, Robert Stevenson. University of Notre Dame, Tech. Rep., 1998.

[2] *Super-Resolution Reconstruction of Images - Static and Dynamic Paradigms.*
Michael   Elad.
http://www.cs.technion.ac.il/~elad/Lectures/
Super-Resolution_All.ppt

[3] *MAP Based Resolution Enhancement of Video Sequences Using a Huber-Markov Random Field Image Prior Model.*
Hu He, Lisimachos P. Kondi. IEEE International Conference on Image Processing, Barcelona, Spain, September 2003, Vol. II, pp. 933-936

[4] *Choice of threshold of the Huber-Markov prior in mapbased video resolution enhancement.*
Hu He, Lisimachos P. Kondi. IEEE Canadian Conference on Electrical and Computer Engineering, Niagara Falls, Canada, May 2004.

[5] *Extraction of High-Resolution Frames from Video Sequences.*
Richard R. Schultz. Robert L. Stevenson. IEEE T. Image Proces., 5(6), pp. 996--1011, 1996.

[6] *A Bayesian Approach to Image Expansion for Improved Definition*
Richard R. Schultz. Robert L. Stevenson. IEEE Transactions on Image Processing, 3(3):233-242, May 1994.

[7] *Aplicaciones de los algoritmos de restauración de imágenes multicanal a problemas de súper – resolución*
Javier Mateos. http://decsai.ugr.es/vip/files/
presentations/2003_mateos.pdf

[8] *Simultaneous Parameter Estimation and Segmentation of Gibbs Random Field using Simulated Annealing*
Sridhar Lakshmanan. Haluk Derin. IEEE Transactions on Pattern Analysis and Machine Intelligence, v.11 n.8, p.799-813, August 1989

[9] *Search Algorithms for Block-Matching in Motion Estimation. Deepak Turaga. Mohamed Alkanhal. Mid-Term project. 18-899. Spring, 1998*

[10] *Block Matching for Object Tracking*
A. Gyaourova. C. Kamath. SC. Cheung. UCRL-TR-200271. 2003.

[11] *Motion Estimation and Hybrid Video Coding Based on Block-M.E.*
Min Wu. ENEE631. Lecture13.

[12] *Markov Random Fields with Applications to M-reps Models.*
Conglin Lu.
http://www.cs.unc.edu/Research/MIDAG/
pubs/presentations/StatsGeomTuts/
LuGeomStats.pdf

[13] *Super-Resolved Surface Reconstruction From Multiple Images*
Peter Cheeseman. Bob Kanefsky. Richard Kraft. John Stutz. Technical Report FIA–94–12

[14] S*úper resolución de imágenes y vídeo*
Rafael Molina.
http://decsai.ugr.es/vip/doctorado/pvd/ T11c.pdf

[15] *Image Inpainting*
Marcelo Bertalmío. Guillermo Sapiro. Vicent Caselles. Coloma Ballester
www.iua.upf.es/~mbertalmio/bertalmi.pdf

[16] *Digital Image Processing.*
R.C. Gonzalez. R.E. Woods. Addison Wesley. 1992

[17] *Processing of Flat and Non-Flat Image Information On Arbitrary Manifolds Using Partial Differential Equations.*
Marcelo Bertalmío.
http://iie.fing.edu.uy/investigacion/grupos/gti/bibli o/doctorado_mbertalmio.pdf

[18] *Image/Video Resolution Enhancement Technique from a Sequence of Low-Resolution Images using Joint MAP Registration Algorithm.*
Hu He. eureka.eng.buffalo.edu/map.pdf