

Generación de Imágenes de Alta Resolución Utilizando Secuencias de Video de Baja Resolución

Federico Cristina

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

Oscar Bria

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

Claudia Russo

crusso@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.*

En ciertos casos, la baja resolución de las imágenes que conforman una filmación de video impide la correcta interpretación visual de la información almacenada en ésta. Un ejemplo típico son los videos obtenidos a partir de cámaras de seguridad.

Existe entonces la necesidad de contar con algún método que permita procesar dicha información para obtener una mejor calidad y un mayor nivel de detalle de éstas.

Esto brinda la posibilidad de realizar una interpretación más confiable de las imágenes, ayudando así a determinar, por ejemplo, los rasgos del rostro de una persona o el número de patente de un automóvil.

Actualmente existen técnicas relacionadas con este tema (denominadas técnicas de Súper Resolución de imágenes) aunque principalmente en el campo teórico. No se cuenta además, con una solución integral presentada como producto integral para su utilización.

Este trabajo presenta resultados preliminares sobre las técnicas de Súper Resolución aplicadas a secuencias de video con la posibilidad de utilizar preprocesamiento de mejora de la calidad a cada imagen individual.

Palabras clave: Super Resolución - Procesamiento de Imágenes - Image Enhancement

Introducción

Al capturar una imagen mediante una cámara digital, o al digitalizar una secuencia de video, se nos presenta el siguiente problema: la información que fotografiamos tiene que ser discretizada y plasmada en píxeles para que pueda ser representada en una computadora. Se pierde entonces tanto información espacial (llevamos una imagen de la vida real a una grilla discreta y finita de píxeles) como información sobre la intensidad de cada píxel (llevamos los niveles de brillo y color a, por ejemplo, una escala de 256 niveles de intensidad en una escala de grises).

Gráficamente, se puede ver un ejemplo simulado en las siguientes imágenes



Figura 1. Imagen a fotografiar

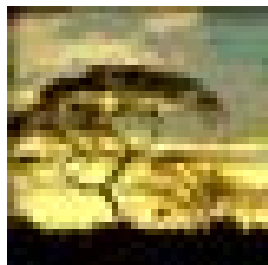


Figura 2. Imagen digitalizada

Sea la Figura 1 una imagen a fotografiar (esto es solo un supuesto ya que, obviamente, ésta ya ha sido digitalizada), notamos como se ha perdido información en ambos sentidos (espacio e intensidad de cada píxel).

Aunque el ejemplo ha sido llevado al extremo; no es inusual que las imágenes de las filmaciones realizadas con cámaras de seguridad sean de calidad pobre. A este problema se le adicionan además otras

complicaciones tales como imágenes con ruido, borroneadas, fuera de foco, etc.

Mejorando una imagen

La forma más simple de obtener una mejora básica en la resolución de una imagen es aplicando a ésta alguna de las técnicas llamadas de interpolación. Las más conocidas son la bicúbica, la bilineal y la técnica del vecino más próximo (nombradas aquí en orden decremental teniendo en cuenta la calidad del resultado obtenido). Aunque dichos métodos presentan una solución rápida, ésta frecuentemente no es suficiente en los peritajes, donde la certidumbre sobre la información observada en la imagen debe ser máxima.

Aunque existen otros métodos que permiten lograr una mejora aún mayor, todavía no son suficientes a los efectos de obtener una optimización significativa.

Aprovechando una secuencia de imágenes

Cuando se cuenta con una secuencia de video donde la información contenida entre un cuadro y el siguiente es aproximadamente la misma (o sea, el movimiento capturado a lo largo de la secuencia de video es relativamente suave); se puede lograr una mejora en el problema detallado anteriormente.

Esto se logra mediante una técnica llamada Súper Resolución de Imágenes [1], [2], [7], [14]; la cual se basa en aprovechar información no redundante de una secuencia de video para obtener como resultado una imagen de mayor resolución.

La Súper Resolución permite minimizar el problema de la discretización y el error de cuantificación. El primer problema se puede resumir en la disyuntiva de determinar en qué píxel colocar cierta parte de la imagen fotografiada; mientras que el segundo se presenta al tener que decidir cuan intenso debe ser dicho píxel, teniendo en

cuenta que se posee un número finito de valores que pueden llegar a asignarse.

La Figura 3 refleja el primer problema. Se puede determinar fácilmente que el punto gris debe ser almacenado en la posición (3,4) de la matriz de píxeles que componen la imagen. Pero, ¿dónde se debe almacenar el punto negro?. (En la práctica, una solución muy utilizada frente a este problema es abarcar más de un píxel para representar el punto. En el ejemplo que acabamos de ver, esto significaría plasmar el punto negro observado en las posiciones (1,1) ; (1,2) ; (2,1) ; (2,2) de la imagen capturada, con una intensidad menor a la que realmente posee dicho punto. De aquí nace el concepto de Point Spread Function (Función de Apertura de un Punto) o PSF, el cual se refiere a cuánto una fracción del objetivo observado incidirá en la zona aledaña a los píxeles que representan dicha fracción en la imagen una vez digitalizada).

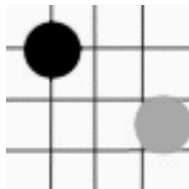


Figura 3. Dos puntos que deben ser plasmados en la imagen

La Figura 4 muestra el problema de la cuantificación. Suponiendo que un píxel (x,y) de una imagen merece un nivel de intensidad de 122.6; dicho píxel deberá ser almacenado con el nivel 123, aunque esto no se corresponda con el valor observado.

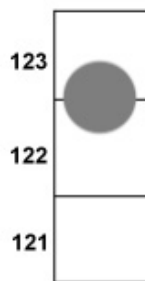


Figura 4. Intensidad lumínica de un píxel

Veamos ahora cómo pueden ayudarnos las técnicas de Súper Resolución con los problemas aquí expuestos.

Básicamente, la solución planteada es generar una media sobre la posición más acertada en la que un píxel representa una fracción de la imagen real. Por lo tanto, si se tienen 3 imágenes similares a la figura 3, donde en la primera el punto negro se almacena en la posición (1,1), en la segunda en la posición (1,2) y la tercera en la posición (2,1); se puede determinar que la solución óptima es almacenar el punto negro en la posición (1,1).

Por otro lado, se puede obtener una media de cada píxel entre las imágenes que conforman la secuencia de video. Así; si se tienen 3 imágenes; en la que el píxel (x,y) en cuestión tiene un valor de intensidad de 123 en la primer imagen, 122 en la segunda y 124 en la tercera; se podrá aseverar que el valor más acertado para éste será de 123.

Aunque de manera simplificada, el fundamento de la Súper Resolución es el arriba explicado. La dificultad inicial está en que generalmente el movimiento de los objetos filmados en la secuencia de video no es suave; o que la cámara misma es la que genera movimiento en el objetivo. El problema en este supuesto es poder determinar qué píxeles de un cuadro corresponden con que píxeles del cuadro anterior.

Es aquí donde aparecen las técnicas de compensación de movimiento [9], [10], [11]; las cuales se utilizan en conjunto con las de Súper Resolución. Gracias a éstas, se puede realizar un mapeo de los píxeles de las imágenes contenidas en un video con respecto a una imagen de referencia de éste.

La figura 5 muestra cómo es la técnica de compensación de movimiento: determina que el píxel (1,1) en la primer imagen se ha movido hasta la posición (3,3) en la segunda. Se obtienen así los vectores de movimiento de cada píxel entre una imagen y la otra.

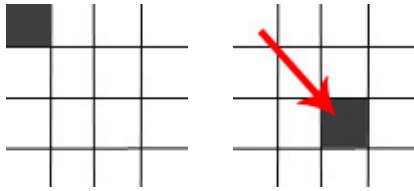


Figura 5. Compensación de movimiento de un píxel (x,y) entre la primer y segunda imagen.

Una complicación más hay que adicionar a este modelo. Generalmente las filmaciones de video contienen imágenes fuera de foco (blurred), con poco contraste, con ruido (periódico y aleatorio), e inclusive imágenes con partes del objetivo de interés tapadas.

La solución a estos problemas es realizar un preprocesamiento individual de cada imagen antes de aplicar compensación de movimiento y Súper Resolución. Dicho preprocesamiento debe ser realizado de acuerdo a las características del video con el cual se va a trabajar. Es por esto que no puede realizarse de manera automática y es más bien un trabajo a medida; aunque de todas formas se intentará aquí poder generar una solución integral que abarque los tres aspectos vistos.

En resumen; las principales etapas en el mejoramiento de la resolución de una imagen partiendo de una secuencia de video son las siguientes:

1. Preprocesamiento individual de cada imagen de acuerdo al problema que éstas presenten (borroneado, ruido, etc)
2. Estimación de los vectores de movimiento entre una imagen referencial perteneciente a la secuencia y el resto de las imágenes que la componen.
3. Aplicar Súper Resolución con las imágenes “limpias” y “compensadas”.

Técnicas estudiadas hasta el momento

Para lograr un resultado óptimo, se están investigando tanto técnicas de preprocesamiento como de compensación de movimiento y Súper Resolución de manera concurrente. Veremos ahora algunos resultados obtenidos:

· Preprocesamiento

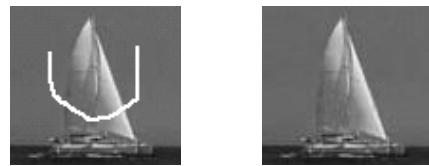


Figura 6a. Inpainting (Desoclusión)



Figura 6b. Inpainting (Desoclusión)

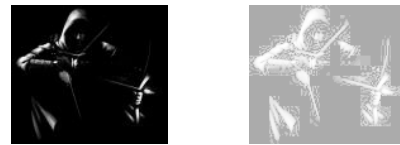


Figura 7. Ecualización del histograma

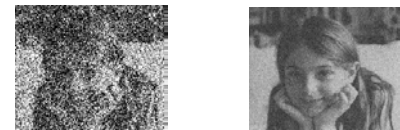


Figura 8. Averaging (Promedio)



Figura 9. Mejora Local



Figura 10. Desenfoque mediante Deconvolución

· Compensación de movimiento



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



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



Figura 11c. Block Matching Motion Compensation (deformation)

· Super Resolución



Figura 12. Interpolación bicúbica vs. técnica de Súper Resolución mediante POCS

Las figuras 6a y 6b muestran la técnica de “inpainting” o “desoclusión” [15], [17] la cual regenera la parte de la imagen faltante a partir de la información aledaña a ésta.

La figura 7 muestra la técnica de ecualización del histograma [16]; la cual ayuda a resaltar las zonas de bajo contraste.

La figura 8 aplica la técnica de promedio, la cual permite eliminar ruido cuando se cuenta con más de una imagen.

La figura 9 se basa en la media y la desviación estándar de la imagen completa y de cada píxel a ser examinado, permitiendo la detección de sutiles cambios en la imagen [16].

La figura 10 permite eliminar el borronado de una imagen mediante la aplicación de una máscara de enfoque, teniendo en cuenta la Point Spread Function.

La figura 11a, 11b y 11c muestran la técnica de Block Matching Motion Compensation (Compensación de Movimiento mediante Comparación de Bloques) en tres situaciones diferentes: movimiento lineal, de rotación y deformación del objeto respectivamente.

La figura 12 muestra la aplicación de la técnica de Súper Resolución basada en Projection Onto Convex Sets (Proyección

hacia Conjuntos Convexos) o POCS para un conjunto de 6 imágenes similares a la superior. La imagen inferior izquierda muestra el resultado mediante la aplicación de interpolación bicúbica; mientras que la inferior derecha muestra el resultado al aplicar Súper Resolución.

Se realizaron procedimientos en casos reales aplicando las técnicas descriptas en el punto anterior para obtener mejoras en la calidad de las fotografías, lográndose la optimización esperada por el usuario.

Futuro de la investigación

La técnica de Súper Resolución mediante POCS genera una marcada diferencia frente a una interpolación bicúbica; sin embargo, no es el método que brinda los mejores resultados. Existe una técnica llamada Maximum A Posteriori (MAP) [5], [6], [13], [18], que se basa en la teoría bayesiana, la cual es notoriamente superior en cuanto a resultados.

En ésta, se realiza un estimado inicial de la imagen mejorada, utilizando el Huber-Markov Random Field (Campo Aleatorio de Huber-Markov) o HMRF como modelo a priori [3], [4], [8], [12]. Luego, dicha imagen se la va optimizando en sucesivas iteraciones.

La mejora es posible mediante la combinación del estimado a priori y un sistema de ecuaciones que relaciona la aproximación a priori con las imágenes de baja resolución que conforman la secuencia de video.

Esto permite resolver el problema en el cual la resolución del sistema de ecuaciones tiene más de una solución. Esta técnica es la que se encuentra actualmente en investigación. Aunque es un método que requiere una gran capacidad de procesamiento; la diferencia en los resultados es muy notoria.

La Súper Resolución mediante MAP utiliza un sistema de compensación de movimiento denominado Hierarchical Subpixel Motion Estimation (Estimación de Movimiento Subpixel Jerárquico) o HSME; puesto que supone que el movimiento de un objeto en una secuencia de video no tiene que ser necesariamente a nivel de píxel.

Se debe distinguir entonces, la diferencia entre el concepto de movimiento a nivel de píxel (el movimiento que realiza un objeto entre dos imágenes; el cual puede ser registrado en uno o más píxeles de distancia entre la primera y la segunda posición) y movimiento a nivel de subpíxel. En este caso el movimiento es muy sutil, lo cual al ser la imagen una matriz discreta no puede reflejar dicho cambio (limitado esto por la resolución de las imágenes), convirtiendo este movimiento a un cambio en las intensidades (de gris por ejemplo) del objeto en cuestión entre las imágenes.

La figura 13 explica de manera simplificada el concepto que se acaba de mencionar. La imagen superior izquierda muestra un punto a ser digitalizado y la imagen superior derecha muestra dicho punto ya capturado en la posición (1,1). Si suponemos que en el siguiente cuadro el punto se ha desplazado como lo indica la imagen inferior izquierda, al plasmar dicho punto en la grilla de píxeles que componen la imagen, se debería hacer almacenándolo en las posiciones (1,1); (1,2) y (2,2); obteniendo así una representación aproximada de la nueva posición del punto; tal como lo indica la imagen inferior derecha. Esta representación será afectada también por otra parte del objetivo capturado; como se puede observar en la figura 14.

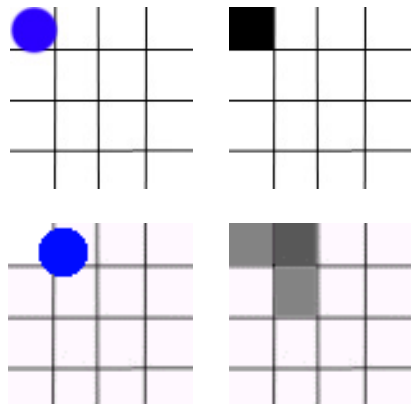


Figura 13. Movimiento Subpixel & PSF

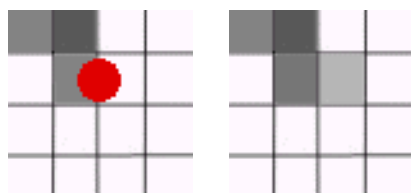


Figura 14. Incidencia de la PSF

Tener en cuenta tanto el movimiento de los objetos entre cuadros a nivel de píxel y subpíxel es vital para lograr un resultado final óptimo.

Resultados esperados

El objetivo inicial de la investigación es poder desarrollar un algoritmo de Súper Resolución mediante MAP aplicando HSME. Se prioriza aquí la calidad de los resultados sobre la performance.

La siguiente etapa abarcaría la integración de la Súper Resolución desarrollada en la etapa anterior con las técnicas de preprocesamiento individual de las imágenes que componen la secuencia de video.

La etapa final de la investigación cubriría el desarrollo de una solución de software integral y optimizada que abarque las técnicas expuestas aquí para tener así un producto completo y accesible a usuarios finales.

Bibliografía

[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] *Súper 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.iaa.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://ie.fing.edu.uy/investigacion/grupos/gti/biblio/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