Digital Image Analysis and Enhancement

S. Battiato, G. DiBlasi, G. Gallo, A. LoGiuoco, M. Madonia, S. Nicotra, F. Stanco

DIPARTIMENTO DI MATEMATICA ED INFORMATICA – UNIVERSITA' DI CATANIA
{battiato, gdiblasi, gallo, logiuoco, madonia, snicotra, stanco}@dmi.unict.it

Abstract. This note reports the main research results of the *Imaging and Graphic Research Group* in the field of image analysis and enhancement. Particular emphasis, reporting also some major details, is given to methods and techniques devoted to unsupervised image analysis and classification, automatic text discrimination, adaptive interpolation and zooming, color and contrast enhancement.

Introduction

This report describes research activities carried out by the *Imaging and Graphic Research Group* working at Dipartimento di Matematica ed Informatica of the University of Catania. Two main topics have been mainly focused: image processing (analysis, enhancement and compression) and pattern recognition (unsupervised classification) both applied to typical machine vision tasks. Research contracts have been achieved by means of "Progetti d'Ateneo", cooperation with STMicroelectronics-Catania and by MIUR project "Trattamento delle Immagini Digitali", cluster 15.

Two main lines of research have been followed: image analysis (medical, sat and aerial) and image enhancement of digital pictures acquired by consumer devices. Of course, the best results in terms of quality enhancement can be obtained by using effectively all the information collected in the analysis phase.

The rest of the paper will describe main results obtained in the various topics; some slightly details will be reported only for two recent techniques regarding respectively aerial image classification and expected color rendition.

Image Analysis

Public and private organizations involved in the acquisition of aerial and satellite images, need suitable analysis techniques to effectively take advantage of the great amount of raw information. The specific expertise derives from participation in European projects and by various partnership with companies operating in the GIS area ([2]) The analysis of aerial images presents some specific characteristics. There exist large databases of these images, produced in the past decades using analogical devices. These images are typically acquired storing only the luminance channel. The analysis, classification and segmentation cannot be performed using classical multichannel techniques already available for satellite images. For these reasons to perform the classification of these images, we introduced some *features* coming from the textural analysis: Laws filtering, morphological analysis and Gabor filtering. These *features* have been the starting point for different classification strategies, partially based on some previous works in the textural analysis and processing ([1],[3],[9]). It has been used a real case study: the analysis of a temporal sequence of aerial images of the Stelvio's natural park.

The first investigated technique is based on the learning of a *membership fuzzy function* relative to different territorial prototypes previously extracted from some test images by an human expert. The *features* used for the classification are the Laws filters. Using Laws filters for gradient evaluation, the overall analysis is not influenced by the luminance variation, obtaining also a good invariance with respect to rotation and translation. The gradient computation allows to manage the micro-structure of the textures present in the neighborhood of each pixel. The final classifier shows *performance* similar to the state of the art for this kind of problem, as can be seen in Figure 1. The robustness of the system is also increased from the fuzzy (not crisp) classification. The algorithm is very efficient whenever it requires a reliable set of prototypes used as reference ([12]).

Together with the fuzzy approach, another technique is actually under development. This method is mainly based on bayesian models (i.e. probabilistic distances derived from Gabor filtering). In particular using a suitable set of directional Gabor filters we obtained a good classification on the same set of images used in previous technique ([13]).

To have a robust classifier for temporal analysis and segmentation on images acquired in different decades, with different sensors and climatic condition, the generalization of these techniques is an important point of our research. Recently, we extended the above strategies to the classification of bad microchip.

We obtained very good results in the field of the discrimination of document images ([6]). By properly using a block partitioning of a digital image (typically acquired by a consumer device) we have designed a processing *pipeline* able to filtering the input data isolating in an effective way the "text" areas, without using classical complex methodologies and techniques such as segmentation, textual analysis, etc. To discriminate the semantic category of an image "text" allows to select a suitable *pipeline* able to outperform the *enhancement* and compression steps. For example this kind of image can be efficiently compressed using an "ad hoc" indexed palette ([11]).

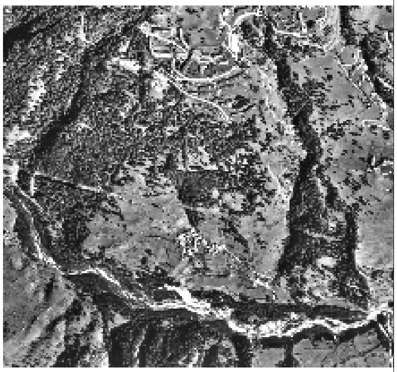
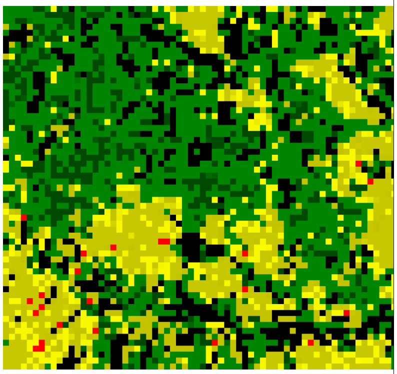
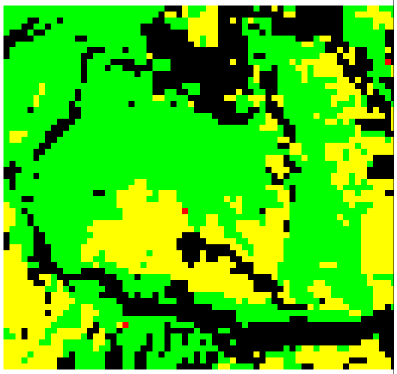
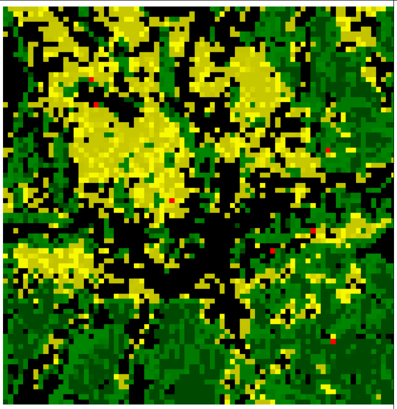
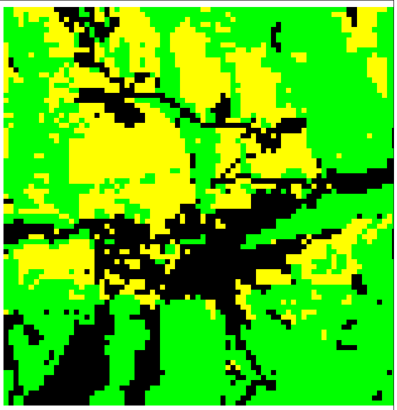
Original Image	***Automatic Classified Image***	***Manually Classified Image***
		
		

Figure 1: Some examples of classified images obtained with our strategy. First column is the original image, second column is the automatic classified image, third column is the manually classified image.

Image Enhancement

Due to the increasing capacity, both in terms of processing power and memory resources, the enhancement of images acquired by consumer devices (Digital Still Cameras, PDA, Mobile phone, etc.) can be designed using advanced ad-hoc techniques. The relative algorithms are implemented in an embedded way, according to the specific HW platform. The enhancement process is devoted to improve the global visual appearance of the input data, working on some specific peculiarities (resolution, visual artefacts, contrast, sharpness, chromaticity, …).

It has been developed an adaptive interpolation algorithm able to outperform the classical zooming algorithm (bicubic, bilinear, etc.) with respect to the overall perceptual quality. The local adaptiveness takes into account magnitude and direction of the local gradient to properly guide the interpolation process ([4]).

In cooperation with the AST (Advanced System Technology) research group of STMicroelectronics of Catania, we have investigated about the performances of several classic and advanced zooming techniques in the specific context of Super-Resolution, reporting a series of comparative studies ([5],[8]).

Among others, it has been also designed an enhancement processing pipeline devoted to mask typical artefacts of images acquired by low-cost imaging devices ([7]) and a more advanced technique able to enhance the global contrast appearance, making use of a local analysis able to detect some "visual significance" areas (e.g. texture, edge, skin, etc.) needed to properly tuning the global contrast setting ([10]).

A new enhancement algorithm semantic based has been recently proposed ([14]). While most enhancement techniques are completely blind to scene content, the proposed technique aims to improve the visual quality of natural scene images by strongly relying on actual, and expected, image colour appearance. The design of the algorithm has been preceded by the collection of reliable statistics on sample images, to be used for the purpose of image classification, allowing at the same time the identification of target colours for the various classes. It is based on a two

steps process: a scene classifier aimed to label each pixel as belonging to a particular chromatic class, followed by an automatic colour correction step with dynamic range and intensity level preserving capabilities. A series of subjective experiments, in which involved people were asked about the perceived quality of the colour corrected images, confirm the effectiveness of the proposed algorithm even when compared with some of commercial available tools.
The algorithm can be viewed as composed of two different processes, image classification and color enhancement. The classification analyzes the input image, and relying on experimentally derived statistics and rules, assigns pixels to a set of fixed semantic classes. The classification, after a further refinement processing, is feed to the enhancement block to drive the adaptive, modulated, class based, color rendition process. The design of the algorithm relies on consistent color statistics for each considered class that have been obtained from a collected sample images database. Figure 2 and 3 show a block based description of the algorithm.

Figure 2. A block description of the various steps employed to generate the classification mask *M*.

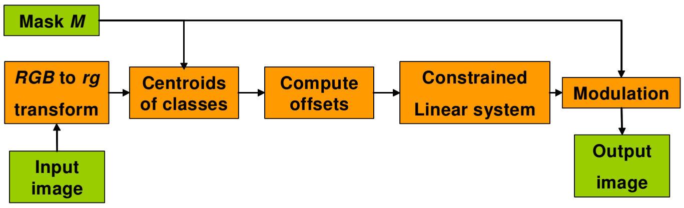


Figure 3. A block description of the steps performed by the color enhancement step. The mask image *M* is used both in the classification and modulation phase.

The process, mainly driven by the collected color statistics, leading from classification of the input image to the final pseudo image mask representing the membership, and the degree of membership of each pixel to the various classes, can be outlined as composed of different steps (see Figure 2).

Punctual classification

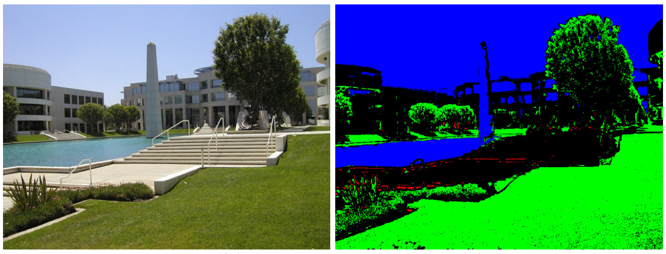
The image is first classified on a per pixel basis using rules that have been easily derived from the collected statistics. In order to avoid dealing with ambiguous values coming from de-saturated and/or low-lit pixels, only pixels satisfying the following condition are considered:

$$(S_k > T_s) \wedge (L_k > T_L) \tag{1}$$

where S_k, L_k are respectively saturation and lightness values for pixel in position k, and T_s, T_L are experimentally fixed thresholds. The assignment to each pixel P_k to the available classes is handled by three mutual exclusive rules:

$$\begin{gathered}\left((H_k \le L_c) \wedge (H_k \ge R_c)\right) \rightarrow P_k \in class_c \\ c \in \{skin, veg, sky\}\end{gathered} \tag{2}$$

For sake of clarity we choose to code with *RGB* triplets the various classes: (255,0,0) for $class_{skin}$, (0,255,0) for $class_{veg}$, and (0,0,255) for $class_{sky}$, coding with (0,0,0) the unclassified pixels. Thus in our convention *R, G* and *B,* are respectively representative of skin, vegetation and sky/sea. Figure 4 shows a sample image and the coded provided by punctual classification step.



(a) **(b)**

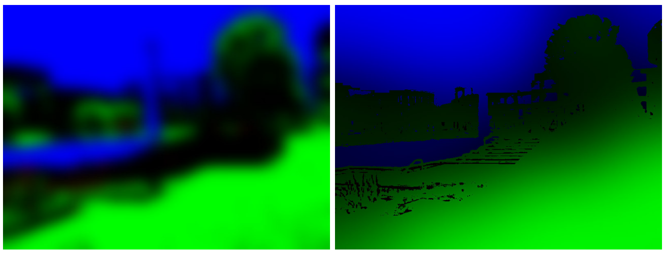
Figure 4. An input image (a), and the output of its punctual classification (b). Blue and green pixels are representative of sky/sea and landscape.

Expansion

Punctual classification is likely to be not always perfect, since several pixels, even if identifiable by visual inspection as belonging to the available classes could not be properly recognized due to high deviation from expected hue values. In order to expand the results coming from the punctual classification step, the mask is subjected to a relevant low pass filtering step. The filtering has been performed by employing a Gaussian kernel, which can be defined by Eq. (3):

$$g(x,y) = e^{\frac{-\left(x^2+y^2\right)}{s^2}} \qquad (3)$$

Experimentally we have found that Eq. (3) needs very high s (space constant) values, thus making filtering in the spatial domain computationally expensive. To avoid this, the filtering is performed on a down-sampled mask image followed by successive up sampling by means of bilinear interpolation. An example of such filtering can be seen in Figure 5 (a). The sampling ratio and the kernel size where chosen to be proportional to input image resolution.



(a) **(b)**

Figure 5. The punctual classification of Figure 3 after the expansion step (left), and the refinement step (right).

Refinement

The refinement step produces the final mask $M=\{c_{k,} w_k\}$, indicating for each pixel in position k the class c_k to which it belongs, and the degree of membership w_k**.** Since the filtering step will cause the results of the punctual classification to overlap (e.g. multiple assignments will be available for the same pixel), a max rule is used to obtain one class and one degree of membership for each pixel.

$$\begin{aligned} c_k &= \left\{class_c : c = \max\left(R_k, G_k, B_k\right)\right\} \\ w_k &= \max(R_k, G_k, B_k) \end{aligned} \qquad (4)$$

Figure 5(b) shows an example of mask.

Adaptive Color Correction

The enhancement is aimed to reduce the distance of colors belonging to the various classes from the target values by means of proper, lightness preserving, color shifting. The mask $M=\{c_{k,} w_k\}$ is used to guide this process, by assigning a class related target to the classified pixels, and by modulating the amount of color correction.

Color targets

For each class (*skin, vegetation, sky-sea*) the targets were obtained by mapping the centroids of the collected statistics on the *rg* (RGB normalized) chromaticity plane . Given an *RGB* color, the mapping on the *rg* plane can be defined as:

$$r = \frac{R}{R+G+B} \qquad g = \frac{G}{R+G+B} \tag{5}$$

The computed color targets for each class c will be indicated as (r_c, g_c).

Shift computation

After converting the input image into the *rg* color space employing (5), the mean value on the color plane of each identified color class is computed as follows:

$$\mu_{rc} = \frac{\sum_k (r_k : c_k = c)}{card_c} \qquad \mu_{gc} = \frac{\sum_k (g_k : c_k = c)}{card_c} \tag{6}$$

with $card_c$ representing the cardinality of class c. For each class, the offset from the target color is defined as:

$$\Delta_{rc} = r_c - \mu_{rc} \qquad \Delta_{gc} = g_c - \mu_{gc} \tag{7}$$

Modulated color enhancement

The color enhancement is carried out by shifting each pixel value $(r_k\ ,\ g_k\ )$ by the computed offset and then converting back in the standard *RGB* color space. The ambiguity, due to the "one to many" mapping, of the inverse of Eq.(5) can be advantageously used to define a lightness preserving, constrained linear system:

$$\begin{cases} \dfrac{R_k^{'}}{R_k^{'} + G_k^{'} + B_k^{'}} = r_k + \Delta_{rc} \\ \dfrac{G_k^{'}}{R_k^{'} + G_k^{'} + B_k^{'}} = g_k + \Delta_{gc} \\ \dfrac{R_k^{'} + G_k^{'} + B_k^{'}}{3} = \dfrac{R_k + G_k + B_k}{3} \end{cases} \tag{8}$$

where (R_k, G_k, B_k) is the input color for pixel k, and (R'_k, G'_k, B'_k) its output value. In order to avoid the appearance of unpleasant artifacts and/or excessive color distortions, the final color correction is modulated by using the computed membership values w_k of the mask M, and two modifiable parameters a and b. The final values $(R_k^{''}, G_k^{''}, B_k^{''})$ are thus defined as follows:

$$R_k^{''} = \frac{aR_k + b\left[w_k R_k^{'} + (1-w_k)R_k\right]}{a+b} \quad (9)$$

$$G_k^{''} = \frac{aG_k + b\left[w_k G_k^{'} + (1-w_k)G_k\right]}{a+b}$$

$$B_k^{''} = \frac{aB_k + b\left[w_k B_k^{'} + (1-w_k)B_k\right]}{a+b}$$

Parameters a and b allow to perform a linear combination between original and color corrected pixel values, while weights w_k decrease or increase the amount of correction depending on the reliability of the classification. This approach allows us to preserve the dynamic range of the classified regions avoiding also a naturalness modification. Two examples of input-output couples relative to the entire process are reported in Figure 6. The Figure also contains an image of the absolute difference between the original and the processed image, that shows how the correction is spatial variant.

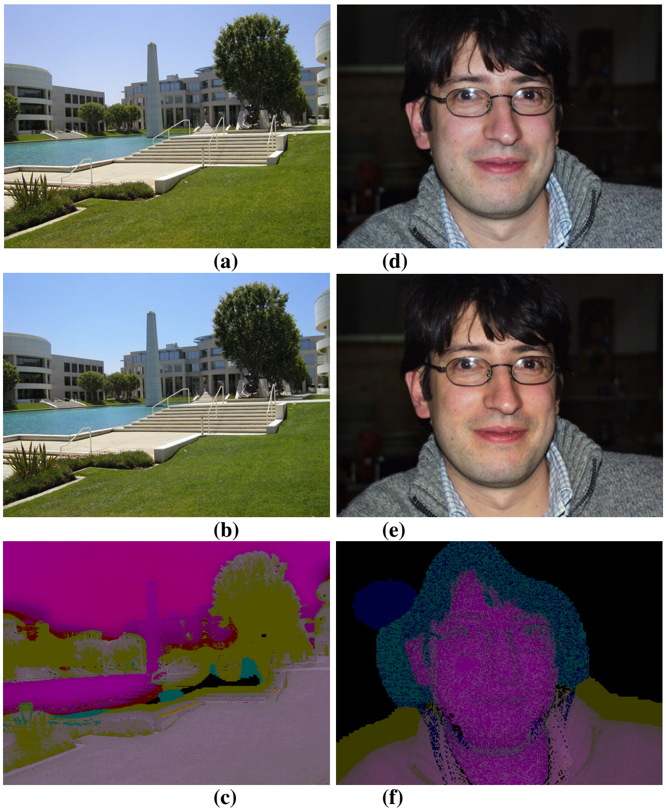


Figure 6. A landscape image (a), its enhanced version (b) and the difference image (c). A portrait image (d) its enhanced version (e) and the difference image (f).

Conclusions and Future Works

In this paper we have briefly summarized the main research results of the *Imaging and Graphic Research Group*, working at Dipartimento di Matematica ed Informatica, University of Catania. Main topics have regarded the investigation of innovative solution for digital image analysis and enhancement. Some detail has been provided about a novel technique able to enhance the colour appearance of digital images.

Future works will include investigation about SVG raster to vector conversion and advanced rate-control techniques for still/video application.

References

[1] S. Battiato, G. Gallo - *Multi-resolution Clustering of Texture Images* – Chapter in Texture Analysis in Machine Vision, Ed. M. Pietikäinen - Series in Machine Perception and Artificial Intelligence - Vol. 40, pp.41-51, World Scientific, October 2000;

[2] G. Gallo, G. Grasso, S. Nicotra, A. Pulvirenti - *Remote Sensed Images Segmentation Through Shape Refinement* – In Proceedings of IEEE ICIAP'01, Int. Conf. on Image Analysis and Processing, 2001 – Palermo, Italy;

[3] S. Battiato, G. Gallo, S. Nicotra – *Glyph Representation of Directional Texture Properties* – 10-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, WSCG 2002 – Plzen, Czech Republic, February 2002 – Published in Journal of WSCG, Vol. 10, No.1-3, pp. 48-54, 2002;

[4] S. Battiato, G. Gallo, F. Stanco – *A Locally-Adaptive Zooming Algorithm for Digital Images* – Image Vision and Computing Journal – Elsevier Science Inc. - Vol. 20, Issue 11, pp. 805-812, September 2002;

[5] S. Battiato, G. Gallo, M. Mancuso, G. Messina, F. Stanco - *Analysis and Characterization of Super-Resolution Reconstruction Methods* – In Proceedings of SPIE Electronic Imaging 2003 - Sensors, Cameras, and Applications for Digital Photography, Vol. 5017B-37 – Santa Clara, CA USA, January 2003;

[6] N. Alessi, S. Battiato, G. Gallo, M. Mancuso, F. Stanco – *Low Level Feature's Set for Text Image Discrimination* – In Proceedings of IEEE-EURASIP Workshop on Non Linear Signal and Image Processing – NSIP 2003 – Grado, Italy - June 2003;

[7] S. Battiato, A. Castorina, M. Guarnera, P. Vivirito – *A Global Enhancement Pipeline for Low-cost Imaging Devices* – IEEE Transactions on Consumer Electronics – Vol. 49, Issue 3, pp. 670-675, August 2003;

[8] S. Battiato, G. Gallo, F. Stanco - *Smart Interpolation by Anisotropic Diffusion* – In Proceedings of IEEE ICIAP'03 Int. Conf. on Image Analysis and Processing – pp. 572-577 - Mantova, Italy – September 2003;

[9] S. Battiato, G. Gallo, S. Nicotra - *Perceptive Visual Texture Classification and Retrieval* – In Proceedings of IEEE ICIAP'03 Int. Conf. on Image Analysis and Processing – pp. 524-529 - Mantova, Italy – September 2003;

[10] S. Battiato, A. Bosco, A. Castorina, G. Messina – *Automatic Image Enhancement by Content Dependent Exposure Correction* – EURASIP Journal on Applied Signal Processing – In press (2004);

[11] S. Battiato, G. Gallo, G. Impoco, F. Stanco – *An Efficient Re-indexing Algorithm for Colour-Mapped Images* – IEEE Transactions on Image Processing – In press (2004);

[12] G. DiBlasi, G. Gallo, S. Moschetto – *Texture Classification in Aerial Photos* - Preprint (2004);

[13] S. Battiato, A. Carini, G. Gallo, M. Madonia – *Gabor filtering for Texture Classification in Geographical Data* – Preprint (2004);

[14] F. Naccari, S. Battiato, A. Bruna, S. Cariolo, A. Castorina – *Natural Scenes Enhancement by Adaptive Color Correction* – In Proceedings of IEEE ISCE 2004 – Int. Symposium on Consumer Electronics, Reading (UK), September 2004.