Low-quality image enhancement using visual attention

Francesca Gasparini, Silvia Corchs, and Raimondo Schettini
Università degli Studi di Milano-Bicocca, DISCo, Via Bicocca degli Arcimboldi 8, 20126 Milano, Italy
E-mail: gasparini@disco.unimib.it

Abstract. Low quality images are often corrupted by artifacts and generally need to be heavily processed to become visually pleasing. We present a modified version of unsharp masking that is able to perform image smoothing, while not only preserving but also enhancing the salient details in images. The premise supporting the work is that biological vision and image reproduction share common principles. The key idea is to process the image locally according to topographic maps obtained from a neurodynamical model of visual attention. In this way, the unsharp masking algorithm becomes local and adaptive, enhancing the edges differently according to human perception. © 2007 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.2721764]

Subject terms: image enhancement; visual attention; unsharp masking; denoise.

Paper 060911LR received Nov. 23, 2006; revised manuscript received Jan. 22, 2007; accepted for publication Jan. 29, 2007; published online Apr. 9, 2007.

1 Introduction

Digital images are often corrupted by artifacts due to noise in the imaging system, digitization, and compression. Smoothing is a widely used technique to obtain more visually pleasing images. Several methods have been proposed in the literature to reduce edge blurring when smoothing is applied. Among the edge sharpening techniques, the unsharp masking approach is widely used to improve the perceptual quality of an image. This is a linear filtering, and although it is easy to implement, it also enhances noises, digitization effects, and blocking artifacts. This often results in visually unpleasing images. Several methods have also been proposed in the literature that focus on smoothing and preserving the details in digital images. In particular, nonlinear techniques have been extensively used. Polynomial filters are frequently used to replace the high-pass filter in the unsharp masking scheme. In Ref. 3, the proposed structure is similar to the conventional unsharp masking, except that the enhancement is allowed only in the direction of the maximal change, and the enhancement parameter is computed as a nonlinear function of the rate of change. A modification of the Laplacian, which is called the order statistic (OS) Laplacian, is considered in Ref. 4. A variable length nonlinear filter that has the capability to both sharpen the edges and smooth out the noise is presented in Ref. 5. In Ref. 6, an adaptive algorithm is introduced so that a sharpening action is performed only in locations where the image exhibits significant dynamics. An adaptive directional filtering is also performed to provide suitable emphasis on the different directional characteristics of the details. However, the solutions cited may still introduce some artifacts due to the amplification of the input disturbances, in particular within the flat areas. They often require weighting parameters, which have to be set on the basis of human intuition or judgment.

In this paper, we present a modified version of the unsharp masking scheme that is able to enhance image quality. This method performs image smoothing, not only preserving but also enhancing the salient details in images. Our algorithm is based on the consideration that there is a strong relationship between biological vision and image rendering. In particular, the image rendering process will be more successful in interpreting the original scene and in applying the appropriate transformations. The key idea is to locally process the image according to topographic maps obtained from a neurodynamical model of visual attention. This overcomes the trade-off between smoothing and sharpening, typical of the traditional approaches.

2 Neurodynamical Model of Visual Attention

In the last decades, several computational attention methodologies have been studied, and they have become a powerful tool in vision systems. Computational neuroscience provides a mathematical framework to study the mechanisms involved in brain functions, such as visual attention mechanisms. It is well known that only part of the visual information of a given scene is processed in full detail, while the remainder is left relatively unprocessed. Therefore, visual attention facilitates the processing of that limited portion of the input associated with the relevant information, and it suppresses the remaining information. The first attention model was proposed by Koch and Ullman.7 Itti and Koch,8 then defined a visual attention system based on saliency maps to predict visually salient features of a scene. Chauvin et al.9 and Gueronnet et al.10 proposed models inspired by the retina and the primary visual cortex cell functionalities. Corchs and Deco11 implemented a neurodynamical model for visual attention, based on evidence from functional, neurophysiological, and psychological findings.12 This visual model consists of interconnected modules that can be related to certain components of the visual cortex. In the application of this paper, we have adopted a reduced version of this model, which is the bottom-up component, given by the module that represents the primary visual cortex. Given an input image, this information enters the visual cortex through this module, named V1. From this theoretical model, we can calculate the neural activities of the V1 neurons, which correspond to the internal representation of the image. This representation is the map of activities of the given input image (Fig. 1). The mathematical formulation and a detailed explanation of the neurodynamical model can be found in Refs. 12 and 13.

3 Low-Quality Image Enhancement Algorithm

In our algorithm, the enhanced image $f_{\text{enh}}(x,y)$ is obtained by subtracting nonsalient high frequencies from the original image $f(x,y)$, producing a smoothing effect, and by adding salient high frequencies, $f_{\text{sal}}(x,y)$, with a consequent edge sharpening.

$$fhS(x,y) = f_{\text{sal}}(x,y) - f_{\text{enh}}(x,y)$$
\begin{equation}
f(x,y) = f(x,y) - \lambda \cdot f_{hNS}(x,y) + \lambda \cdot f_{hS}(x,y).
\end{equation}

In Eq. (1), salient high frequencies \( f_{hS}(x,y) \) are obtained by weighting all of the image high frequencies through a topographic map corresponding to visually salient regions, obtained by the neurodynamical model of visual attention described in Sec. 2. This enhances the edges differently without enhancing the perceived noise. On the other hand, nonsalient high frequencies \( f_{hNS}(x,y) \) are simply obtained by weighting all of the high frequencies \( f_{h}(x,y) \) through the complement of the same topographic map, causing an adaptive smoothing. In our implementation, the high frequencies \( f_{h}(x,y) \) are evaluated using the Laplacian, which is the simplest isotropic derivative operator. It should be noted that our enhancement method is independent of the high-pass filter used.

4 Experimental Results

We have tested our algorithm on a data set composed of 300 color images of different sizes (from 120 x 160 to 2272 x 1704 pixels), resolutions, and quality. These images are mostly landscapes, portraits of people, and indoor groups of people and were downloaded from personal Web pages or acquired using various digital cameras and scanners. The quality of these images in terms of JPEG compression, noise, dynamic range, and so on varied and was in most cases unknown. The edge enhancement algorithm works only on the intensity channel to avoid possible color distortion due to saturation. We have compared our algorithm with the traditional unsharp masking applied to the original image previously smoothed by a Gaussian filter. The sigma of this filter varies according to the image size, and it is chosen to produce a smoothing equivalent to the remaining 9% of the cases, the two methods have been judged equivalent. The enhanced images were always judged better than the not-processed original images. More examples are available at www.ivl.disco.unimib.it/Activities/Edge%20sharpening.htm

5 Conclusions

This paper presents a modified version of the unsharp masking method that makes it possible to reduce the sharpening sensitivity to noise and also due to digitization effects and compression. Moreover, being based on a neurodynamical model of human visual attention, the processed images appear more natural and pleasing to observers. The preliminary experimental results on a data set of about 300 images of various subjects with various types of degradations demonstrate that our method has good performance with respect to the traditional unsharp masking combined with smoothing. We plan to investigate further the relationship among user preferences, image content, image defects.
and artifacts, and our image enhancement based on the Corchs-Deco visual attention model. To this end, a completely new data set is under construction in which the images are degraded in a completely controlled manner. Such a study will also include the influence of the adopted visual attention model in the proposed approach. Different visual attention models generate maps that may be visually different. Our preliminary study indicates the suitability of the Corchs-Deco model because it identifies regions that present strong edges and that correspond to high contrast areas. However, the model is somewhat expensive from the computational point of view. Other models such as Ref. 8 should be implemented therefore and extensively compared with ours in terms of both effectiveness and efficiency. A comparison with more sophisticated enhancement methods (e.g., Refs. 5 and 6) would also be interesting.

References