Available online at www.sciencedirect.com
SCIENCE DIRECT®

Appl. Comput. Harmon. Anal. 19 (2005) 340–358

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)

M. Elad [a,*], J.-L. Starck [b], P. Querre [b], D.L. Donoho [c]

[a] *Computer Science Department, The Technion—Israel Institute of Technology, Haifa 32000, Israel*
[b] *CEA-Saclay, DAPNIA/SEDI-SAP, Service d'Astrophysique, F-91191 Gif sur Yvette, France*
[c] *Department of Statistics, Stanford University, Sequoia Hall, Stanford, CA 94305, USA*

Received 27 October 2004; revised 6 March 2005; accepted 9 March 2005

Available online 15 August 2005

Communicated by Charles K. Chui

Abstract

This paper describes a novel inpainting algorithm that is capable of filling in holes in overlapping texture and cartoon image layers. This algorithm is a direct extension of a recently developed sparse-representation-based image decomposition method called MCA (morphological component analysis), designed for the separation of linearly combined texture and cartoon layers in a given image (see [J.-L. Starck, M. Elad, D.L. Donoho, Image decomposition via the combination of sparse representations and a variational approach, IEEE Trans. Image Process. (2004), in press] and [J.-L. Starck, M. Elad, D.L. Donoho, Redundant multiscale transforms and their application for morphological component analysis, Adv. Imag. Electron Phys. (2004) 132]). In this extension, missing pixels fit naturally into the separation framework, producing separate layers as a by-product of the inpainting process. As opposed to the inpainting system proposed by Bertalmio et al., where image decomposition and filling-in stages were separated as two blocks in an overall system, the new approach considers separation, hole-filling, and denoising as one unified task. We demonstrate the performance of the new approach via several examples.

Keywords: Basis pursuit; Total variation; Sparse representation; Cartoon; Texture; Inpainting

* Corresponding author.
E-mail addresses: elad@cs.technion.ac.il (M. Elad), jstarck@cea.fr (J.-L. Starck), donoho@stat.stanford.edu (D.L. Donoho).

1063-5203/$ – see front matter
doi:10.1016/j.acha.2005.03.005

1. Introduction

Filling-in 'holes' in images is an interesting and important inverse problem with many applications. Removal of scratches in old photos, removal of overlaid text or graphics, filling-in missing blocks in unreliably transmitted images, scaling-up images, predicting values in images for better compression, and more, are all manifestations of the above problem. In recent years this topic attracted much interest, and many contributions have been proposed for the solution of this interpolation task. Common to these many techniques is the understanding that classic interpolation methods (such as polynomial-based approaches) are not satisfying; indeed nonlinear strategies and local adaptivity seem crucial.

Among the numerous approaches to fill in holes in images, variational methods are very attractive; these were pioneered by Guillermo Sapiro and his collaborators [6,20,21], and followed by Chan and Shen [7]. These techniques were coined *Inpainting* as a reminder of the recovery process museums experts do for old and deteriorating artwork. In their work, Sapiro et al. motivate the filling-in algorithms by geometrical considerations: one should fill in by a smooth continuation of isophotes. This principle leads to one or another nonlinear partial differential equation (PDE) model, propagating information from the boundaries of the holes while guaranteeing smoothness of some sort. In a series of publications, the geometric principle has been implemented through several different PDEs, aiming to get the most convincing outcome.

The variational approach has been shown to perform well on piecewise smooth images. Here and below we call such images *cartoons*, and think of them as carrying only geometric information. Real images also contain textured regions, and variational methods generally fail in such settings. On the other hand, local statistical analysis and prediction have been shown to perform well at filling in texture content [3,13,29].

Of course real images contain both geometry and texture; they demand approaches that work for images containing both cartoon and texture layers. In addition, approaches based on image segmentation—labeling each pixel as either cartoon or texture—are to be avoided, since some areas in the image contain contributions from both layers. Instead, a method of additively decomposing the image into layers would be preferred, allowing a combination of layer-specific methods for filling in.

This motivated the approach in [2]. Building on the image decomposition method by Vese, Osher, and others [1,28], the image was separated into cartoon and texture images. The inpainting was done separately in each layer, and the completed layers were superposed to form the output image. The layer decomposition, a central component in this approach, was built on variational grounds as well, extending the notion of total-variation [23], based on a recent model for texture images by Meyer [22]. An interesting feature of this overall system is that even if the image decomposition is not fully successful, the final inpainting results can be still quite good, since the expected failures are in areas where the assignment to cartoon/texture contents is mixed, where both inpainting techniques perform rather well.

In previous papers we presented an alternative approach to layer decomposition, optimizing the sparsity of each layer's representation [25,26]. The central idea is to use two adapted dictionaries, one adapted to represent textures, and the other to represent cartoons. The dictionaries are mutually incoherent; each leads to sparse representations for its intended content type, while yielding nonsparse representations on the other content type. These are amalgamated into one combined dictionary, and the basis-pursuit denoising (BPDN) algorithm [8] is relied upon for proper separation, as it seeks the combined sparsest solution, which should agree with the sparse representation of each layer separately. This algorithm was shown to perform well, and was further improved by imposing total-variation (TV) regularization as an