

# Image Completion by Diffusion Maps and Spectral Relaxation

Shai Gepshtein and Yosi Keller

**Abstract**—We present a framework for image inpainting that utilizes the diffusion framework approach to spectral dimensionality reduction. We show that on formulating the inpainting problem in the embedding domain, the domain to be inpainted is smoother in general, particularly for the textured images. Thus, the textured images can be inpainted through simple exemplar-based and variational methods. We discuss the properties of the induced smoothness and relate it to the underlying assumptions used in contemporary inpainting schemes. As the diffusion embedding is nonlinear and noninvertible, we propose a novel computational approach to approximate the inverse mapping from the inpainted embedding space to the image domain. We formulate the mapping as a discrete optimization problem, solved through spectral relaxation. The effectiveness of the presented method is exemplified by inpainting real images, where it is shown to compare favorably with contemporary state-of-the-art schemes.

**Index Terms**—Image inpainting, texture synthesis.

## I. INTRODUCTION

**I**MAGE inpainting algorithms aim to fill the missing data of an image or a video sequence in a visually plausible way, such that the insertion is not easily detectable by a common unsuspecting viewer. It was initially proposed as a mean for restoring old and scratched pictures or artworks [1]. With the proliferation of digital cameras in general, cell phones equipped with cameras in particular, and the increase in consumer-level computational power, digital image manipulation is becoming ubiquitous [2], [3]. Thus, one may remove an undesired object in the scene, and fill the missing part with its background as if the undesired object was never there. This paves the way to new approaches to image editing [4] and adaptive resizing [2].

The image completion problem lies at the intersection of computer graphics, image, and signal processing. Formally, given a corrupted image  $I$  and a “hole” region mask  $H$  marking the unknown area, the goal of image inpainting is to fill in  $H$  to form a visually plausible image  $\hat{I}$ . This is depicted in Fig. 1(a). The problem has attracted significant research effort, and existing methods can be categorized as being either exemplar-based or variational schemes. Variational schemes

Manuscript received December 29, 2011; revised July 17, 2012; accepted October 29, 2012. Date of publication January 11, 2013; date of current version May 24, 2013. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Theo Gevers.

The authors are with the Faculty of Engineering, Bar Ilan University, Ramat Gan 52900, Israel (e-mail: shaige@ gmail.com; yosi.keller@ gmail.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2013.2237916

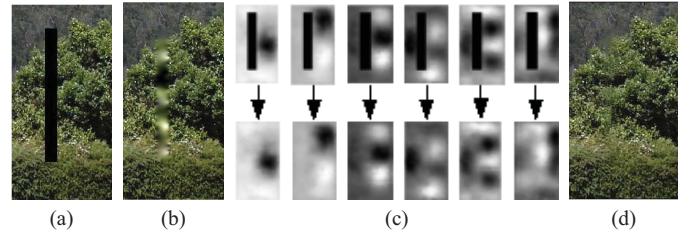


Fig. 1. Image inpainting. (a) Source image to be inpainted. (b) Source image inpainted by an isotropic heat equation. (c) Leading eigenvectors inpainted using an isotropic heat equation. (d) Image inpainted through computing the inverse diffusion mapping of (c).

[4]–[9] formulate the inpainting as the minimization of a variational functional that encodes spatial smoothness constraints. The minimization of such functionals results in the solution of linear or nonlinear heat equations. Thus, these schemes assume that the inpainted images contain piecewise smooth spatial manifolds. Exemplar-based schemes [10]–[14] implicitly assume that  $H$  can be inpainted by replicating image patches given in a reference set  $\bar{H}$ , that might consist of the known parts of the input image  $I = \bar{H} \cup H$ , or a given set of images [15], [16]. Such schemes were proven to provide efficient solutions, but lack the global optimality of the variational approaches. A combination of these two classes of methods were also proposed utilizing both global geometric completion and texture synthesis [14], [17], [18].

In this work, we propose a new image inpainting framework where the completion algorithm is executed in a diffusion map feature space [19]. Each pixel  $x_{ij} \in I$  in the reference set  $\bar{H}$  is represented by a patch  $p_{ij} \in \mathbb{R}^D$ , thus forming the reference set  $\{p_{ij}\}$ , where  $D$  is typically  $5 \times 5$  or  $7 \times 7$ . To analyze this high dimensional data set, we use diffusion maps for dimensionality reduction. Thus, we implicitly assume that the unknown image patches  $p_{ij} \in H$  and the patches in the reference set  $p_{ij} \in \bar{H}$  belong to the same low dimensional manifold, parameterized by the diffusion kernel. The dimensionality reduction is feasible due to the low intrinsic dimensionality of the image patches, manifested by their local correlations [20]. Using the embeddings, patches are mapped into a low dimensional space

$$\bar{H} \mapsto \Psi(\bar{H}), \Psi \in \mathbb{R}^d, d \ll D \quad (1)$$

This is shown in Fig. 1, where a textured image contains a rectangular hole  $H$  to be filled. Applying a variational inpainting scheme might result in excessive smoothing [Fig. 1(b)]. However, using diffusion maps to reduce the dimensionality

of the data, the resulting embedding (eigenvectors) shown in Fig. 1(c), is smooth. Due to this smoothness, the holes in the “images” of the eigenvectors can be inpainted using simple interpolation methods. As the Diffusion embedding is nonlinear, and noninvertible, the missing data in the image domain cannot be computed directly. For that, we derive a novel approximate embedding inversion scheme that assigns an image patch  $\mathbf{p}_{ij} \in \overline{H}$  to each inpainted point  $\mathbf{y}_{ij} \in \{\psi_{k_1}(\mathbf{x}_{i,j}) \dots \psi_{k_m}(\mathbf{x}_{i,j})\}^T$  in the embedding domain. The assignment problem is solved via spectral relaxation [21] and induced spatial regularization.

Thus, we provide two core contributions: First, a general framework for data completion by deriving smooth representations via Diffusion embedding. In particular, it provides a unified scheme to inpainting textured and textureless images. Second, we derive a scheme for approximating the inverse of the Diffusion embedding and map the interpolated embeddings back to the data/image domain, by formulating the inverse-mapping as an assignment problem, that optimizes a global smoothness constraint. The solution to the resulting combinatorial problem is NP-hard, and is efficiently approximated by spectral relaxation.

The framework we present is general in nature and can be applied to diverse data sources of interest, such as audio signals and tabular data. Yet, in this work we chose to concentrate on the image inpainting problem that provides an intuitive testbed for data embedding and a baseline of contemporary state-of-the-art work to compare against.

This paper is organized as follows: we start by reviewing previous work on image inpainting in Section II-A and recalling the diffusion framework in Section II-B. The proposed image inpainting schemes is introduced and discussed in Section III. It is experimentally verified and compared with existing approaches in Section IV. Concluding remarks are discussed in Section V.

## II. BACKGROUND

In this section we survey contemporary results in image inpainting (Section II-A) and provide background on the Diffusion Maps (Section II-B) that is the main computational tool used in our work.

### A. Image Inpainting

This study of image inpainting lies at the intersection of image processing and computer graphics. Thus, the fundamentals of many contemporary inpainting schemes can be traced back to the seminal contributions of Efros and Leung [10] (exemplar-based) and Bertalmio [5] *et al.* (variational schemes).

Exemplar based techniques proved successful in large-scale image completion tasks (note the *Bungee* example in Fig. 13). Such methods fill the missing pixels by copying source patches from the observed part of the image to produce plausible visual results. Texture synthesis by nonparametric sampling was first introduced by Efros and Leung [10] that proposed a greedy scheme that operates on the pixels of the boundary  $\partial H$ , and successively fills the hole towards its center. The most similar patch (in terms of the  $L_2$  norm) in  $I$  is copied as the

predicted new value. Such schemes are sensitive to the filling order, and might propagate errors of wrongly selected filling pixels, leading to visual inconsistencies. Igely and Pereira [14] proposed to iteratively inpaint synthesized textures by filling the unknown image region with texture synthesized from a second image. The texture synthesis is computed by adapting the histogram of an input noise signal to match that of the target image.

Criminisi *et al.* [11] extended the previous scheme by introducing priority ordering to derive the order in which the pixels are synthesized. The ordering is given by two terms. The first quantifies the strength of the isophotes hitting the hole’s border, thus encouraging linear structures with high strength of isophotes to be synthesized first. The second encodes the reliability of the information in the pixel’s vicinity. Therefore, patches that have more of their neighboring pixels already filled, will be filled first. The priority assignment is a product of the two terms computed at each pixel and updated in each iteration of the algorithm.

A globally optimal formulation of exemplar based inpainting was suggested by Wexler *et al.* [12], where the objective function adds a constraint that missing pixels have to be consistent with all the surrounding patches forming them. In order to reach a global optimum, the inpainting is iterated until the update converges. This approach utilizes a multiscale formulation by initiating the iterative completion at the coarsest level of the pyramid, and propagating the result upward. This allows improved global consistency and speeds up the convergence. The authors extended this approach [12] by introducing spatial and temporal constraints for video inpainting via a variational formulation.

Another iterative exemplar based scheme was suggested by Drori *et al.* [13], where the unknown region is approximated by classifying the pixels to an underlying structure that agrees with other parts of the image. The approximated region is then augmented with details extracted from high confidence regions. The scheme is iterated in different scales until convergence.

The computational bottleneck of exemplar-based schemes is the search for the patch most similar to the one being inpainted, as this search is repeated at each iteration. Barnes *et al.* [3] proposed an efficient approach to K-nn search of image patches, denoted as Patchmatch that provides an order of magnitude performance improvement over previous state-of-the-art schemes such as KD-trees. Their approach is based on random initiation of the patch matches followed by the propagation of patches in the vicinity of well matched patches. It was successfully applied to applications such as image inpainting, retargeting and reshuffling.

The focal point of variational schemes is to formulate the inpainting problem as the minimization of variational functionals. Using the calculus of variations, the minimization of such functionals boils down to the solution of a nonlinear PDEs [5], where the pixels on the boundary  $\partial H$ , are used as boundary conditions. Thus, image information is propagated from  $\partial H$  into the hole  $H$ . The propagation is governed by the chosen functional and the resulting PDE, where nonlinear anisotropic diffusion is preferred to isotropic diffusion to avoid

over-smoothing. Some schemes [6], [22] propose to directly choose a PDE with desirable inpainting properties, without deriving it by the minimization of a functional and are thus nonvariational. Such schemes were shown to be beneficial when inpainting piecewise smooth images or when the hole  $H$  is relatively narrow. Yet, they might fail when inpainting the textured parts of an image.

An efficient image manipulation and inpainting variational scheme was proposed by Pérez *et al.* [4], where the image inpainting and cloning problems are formulated as a Poisson PDE by setting the PDE's boundary condition based on the boundary of the inpainted region  $\partial H$ , and the PDE in the inpainted (or cloned area) is given by the Laplacian of the source image. One of the first schemes utilizing the propagation approach was suggested by Masnou *et al.* [23], that inpaint by joining points of level lines arriving at the boundary of  $H$  with geodesic curves. The edges recovered by this approach are smooth and continuous at the boundary of the hole, and it shows good results for inpainting thin holes.

Shen derived variational image inpainting schemes [9] based on a Bayesian formulation, where the image priors are based on the Bounded Variations (BV) image model and the Mumford–Shah functional. A combination of variational and exemplar based methods was proposed by Bertalmio [18] and Bugeau [17]. The algorithm combines texture synthesis and geometric methods for image inpainting through structure–texture decomposition. Bugeau [17] applied PDE-based techniques to the structure image that is smoother than the original one, and encodes the high frequencies of the image.

Tschumperle and Deriche [22] proposed a general framework for vector-valued image regularization, that is based on variational methods and PDEs. They derive local filters based on anisotropic diffusion. The approach is shown to be applicable to a class of image processing tasks such as image magnification, restoration and inpainting. An image inpainting scheme that combines exemplar-based and PDE-based approaches was suggested by Le Meur *et al.* [24]. The structure tensor is used to set the filling order in the isophote direction, while the optimal image patch used to inpaint the hole, is found via template matching.

This work of Xue and Dai [25] is of particular interest, as it presents a manifold learning approach to inpaint large-scale color image regions with varying colors. The RGB color features of the inpainted image are embedded using Isomap and the inpainting is conducted in the resulting embedding space using Poisson editing.

An inpainting scheme based on discrete optimization was suggested by Komodakis *et al.* [26], where a discrete variational-like functional is iteratively minimized using priority-belief propagation. The scheme solves a global discrete optimization over all possible patches, using dynamic label-pruning to significantly reduce the number of labels. A different class of recent results is based on sparse reconstruction [16], [27], where an image is represented as a set of image patches that are processed as a set of high dimensional samples for which a sparse dictionary is learnt. The missing data  $H$  is extrapolated by linear combinations of the dictionary atoms, computed via  $L_1$  minimization.

Our work relates to both variational and sparse reconstruction based approaches, as these assume that the inpainted image belongs to a low-dimensional space, being either the space of smooth images (as in the variational approaches) or in the space spanned by a sparse dictionary. Our scheme utilizes a different notion of low dimensionality and smoothness that relates to the Diffusion Framework [19]. By utilizing application-specific affinity measures and kernels (Section III-C) we *induce* a chosen smoothness over the image manifold. Thus, simple variational approaches can be applied to inpaint textured images. These different notions of smoothness are further discussed in Section III-D.

### B. Diffusion Maps

In this Section we survey the diffusion framework, where a broader view can be found in [19] and [20]. Given a set  $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  of  $n$  data points, such that  $\mathbf{x}_i \in \mathbb{R}^D$ . Our goal is to compute a set  $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_n\}$ , where  $\mathbf{y}_i \in \mathbb{R}^d$ , such that  $\mathbf{y}_i$  represent  $\mathbf{x}_i$ . Such representation forms a low dimensional mapping of the data points if  $d \ll D$ .

The core of the Diffusion framework is to represent the input dataset  $X$  by an undirected graph  $G = (V, E)$  with nodes  $V = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ , and edges that quantify the affinity between two “close” points  $\mathbf{x}_i$  and  $\mathbf{x}_j$ , such that  $E \in w_{i,j}$ . The affinity matrix  $W = \{w_{i,j}\}$  is required to be symmetric, nonnegative and is commonly computed using an RBF kernel

$$w_{i,j} = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/\sigma^2) = \exp(-d_{ij}^2/\sigma^2) \quad (2)$$

where  $\sigma > 0$  is a scale parameter.  $\sigma$  is a measure of closeness that may prune edges from the graph such that two points  $\mathbf{x}_i$  and  $\mathbf{x}_j$  will have nonzero affinity  $w_{i,j}$  if their distance  $d_{ij} < 3\sigma$ . This weight is related to the heat kernel and defines the nearest neighbor structures of the graph.

The computation of  $w_{i,j}$  is application-specific. In images, the distance  $d_{ij}$  is chosen to encode different image attributes such as color spaces (RGB, LAB), texture via texture descriptors (such as LBP), and low-level image structures (edges). This is further discussed in the context of our work in Section III-C. Coifman and Lafon [19] proved that any weight of the form  $h(\|\mathbf{x}_i - \mathbf{x}_j\|)$  allows to approximate the heat kernel if  $h$  decays sufficiently fast at infinity, and this justifies the use of the RBF kernel in Eq. 2. The core of the diffusion framework is to induce a random walk on the data set  $X$  by normalizing the weights to a Markov matrix

$$\mathbf{A} = \mathbf{D}^{-1} \mathbf{W} \quad (3)$$

where  $d_{ii} = \sum_j w_{i,j}$  is the degree of node  $\mathbf{x}_i$ .  $\mathbf{A}$  is row stochastic and can be viewed as a random walk process as  $a_{i,j} \geq 0$  and  $\sum_j a_{i,j} = 1$ . The term  $a_{i,j}$  represents the probability to jump from  $\mathbf{x}_i$  to  $\mathbf{x}_j$  in a single time step. The corresponding matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  represents the transitions of this Markov chain in a single step

$$a_{ik}^t = Pr(x(t) = \mathbf{x}_k | x(0) = \mathbf{x}_i). \quad (4)$$

$\mathbf{A}$  encodes a time-homogeneous Markov chain, and taking powers of this matrix amounts to running the chain forward in time.  $\mathbf{A}^t$  represents the probabilities of transition from  $\mathbf{x}_i$

to  $\mathbf{x}_j$  in  $t$  time steps. If the graph is connected, the stationary distribution that satisfies the equation  $\phi_0^T \mathbf{A} = \phi_0^T$  is also the left eigenvector of the transition matrix  $\mathbf{A}$  associated with the eigenvalue  $\lambda_0 = 1$ .

The observation of the weights as probabilities of a random walk process paves the way for a spectral decomposition scheme. The eigendecomposition of the transition matrix with respect to the Markovian time variable  $t$  yields

$$\mathbf{A}^t(x_i, x_j) = \sum_{l \geq 0} \lambda_l^t \psi_l(\mathbf{x}_i) \phi_l(\mathbf{x}_j) \quad (5)$$

where  $\{\lambda_l\}$  is the sequence of eigenvalues of  $\mathbf{A}$  (with  $\lambda_0 = 1$ ) and  $\{\phi_l\}$  and  $\{\psi_l\}$  are the corresponding biorthogonal left and right eigenvectors. Due to the spectrum decay, only a few terms are needed to achieve a given relative accuracy in Eq. 5.

The diffusion distance quantifies the similarity between two points  $\mathbf{x}_i$  and  $\mathbf{x}_j$  by their probability distributions

$$D_t^2(\mathbf{x}_i, \mathbf{x}_j) = \sum_{k \in \Omega} \frac{(a_{ik}^t - a_{jk}^t)^2}{\phi_0(k)}. \quad (6)$$

This quantity is a weighted  $L_2$  distance between the conditional probabilities  $a_{ik}^t$  and  $a_{jk}^t$ , that can be thought of as features measuring the interaction of the nodes  $\mathbf{x}_i$  and  $\mathbf{x}_j$  with the rest of the graph. The relation between the diffusion distance and the eigenvectors is given by

$$D_t^2(\mathbf{x}_i, \mathbf{x}_j) = \sum_{l \geq 1} \lambda_l^{2t} (\psi_l(\mathbf{x}_i) - \psi_l(\mathbf{x}_j))^2. \quad (7)$$

This identity implies that the right eigenvectors can be used to compute the diffusion distance, where the different eigenvectors are weighted by the corresponding eigenvalues  $\{\lambda_l\}$ , and only a few terms are needed to achieve a given relative accuracy in Eq. 7 due to the spectrum decay.

Hence, the right eigenvectors  $\{\psi_l\}$  can be used as a new set of coordinates for the set  $X$ , such that the Euclidean distance between these Diffusion coordinates  $\{\psi_l\}$  approximates the diffusion distance in Eq. 7. Let  $d(t)$  be the number of terms retained, the *diffusion map* (embedding) is given by

$$\Psi_t : \mathbf{x} \mapsto \left( \lambda_1^t \psi_1(\mathbf{x}), \lambda_2^t \psi_2(\mathbf{x}), \dots, \lambda_{d(t)}^t \psi_{d(t)}(\mathbf{x}) \right)^T. \quad (8)$$

Using the diffusion map, we represent a graph of any generic data set as a cloud of points in a Euclidean space, where the chosen affinity measure allows to quantify a particular, application-specific distance measure.

### III. IMAGE COMPLETION

Let  $I$  be an image consisting of an unknown part  $H$  and its complementary  $\overline{H}$ , and let  $\partial H$  be the boundary of  $H$ . We aim to inpaint  $H$  by extending the boundary values  $I(\partial H)$  into  $H$ , as depicted in Fig. 2.

The core of our approach is to represent the image in the Diffusion space discussed in Section II-B, and inpaint the embedding, instead of inpainting the pixel values as in previous work. By choosing an appropriate similarity measure between the pixels  $\mathbf{x}_i \in \overline{H}$ , the embedding can be made smooth, thus allowing simple inpainting in the Diffusion domain. The spatial smoothness of the embedding

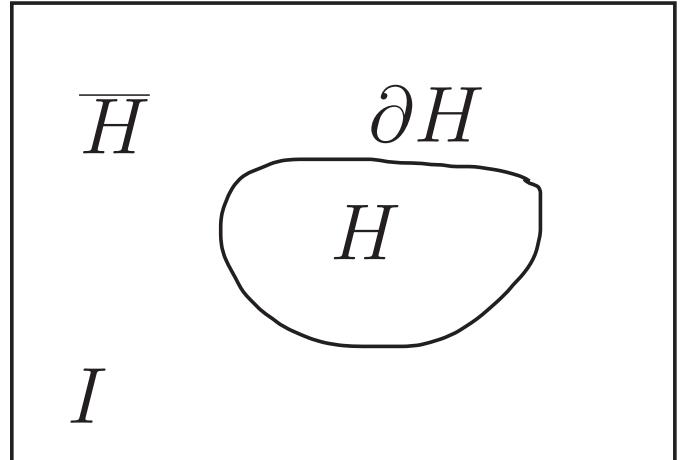


Fig. 2. Image completion problem. We are given an image  $I$ , the unknown part  $H$ , and its complementary  $\overline{H}$ , such that  $I = H \cup \overline{H}$ .  $\partial H$  is the boundary of  $H$ .

vectors stems from them being the eigenvectors of a Sturm-Liouville operator, where similar to the Fourier domain, the leading eigenvector is the smoothest, constant vector, and the following eigenvectors become successively oscillatory, analogue to the higher frequencies in the Fourier domain. This is exemplified in Fig. 1(c).

$\overline{H}$ , the known part of the image, is used as a learning set to compute the embedding of the image manifold. In general,  $\overline{H}$  can be any set of image features, for instance,  $\overline{H}$  can be extracted from a video sequence [12] or a large dataset of images [15]. Following Section II-B, each pixel  $\mathbf{x}_{i,j} \in \overline{H}$  is mapped to a  $d$ -dimensional Diffusion embedding space

$$\Psi : \mathbf{x}_{i,j} \mapsto (\lambda_1 \psi_1(\mathbf{x}_{i,j}), \lambda_2 \psi_2(\mathbf{x}_{i,j}), \dots, \lambda_d \psi_d(\mathbf{x}_{i,j}))^T. \quad (9)$$

This mapping provides  $d$  eigenvectors  $\{\psi_i(\mathbf{x}_{i,j})\}_1^d$  forming a  $d$ -dimensional manifold where “similar” pixels in the image domain are mapped to close locations on the manifold. The notion of similarity is set by the particular choice of the feature space used to compute the embedding  $\Psi$ , and it is common to represent a pixel  $\mathbf{x}_{i,j} \in I$  by an image patch  $\mathbf{p}_{i,j} \in I$ , centered at  $(i, j)$ . The patch-based representation allows to characterize the image texture as discussed in Section III-C. The dimensionality of the embedding space, (the number of eigenvectors used), is related to the intrinsic dimensionality of the image.

The smoothness of the resulting embeddings is exemplified in Fig. 3, where a textured image is embedded using the LBP texture descriptor [28] as an affinity measure that quantifies texture similarity. Thus, although the image is discontinuous in the image domain, patches with similar texture are closely mapped.

Given the embedding  $\Psi \in \mathbb{R}^d$  of the set  $\overline{H}$  we aim to inpaint/extrapolate  $\Psi$  onto  $H$ . For that, one can apply any inpainting/interpolation scheme to the embedding vectors. To exemplify the validity of our approach, we chose to use a simple linear isotropic diffusion

$$\Delta \phi_i = 0, \phi_i(\partial H) = \psi_i(\partial H). i = 1..d. \quad (10)$$

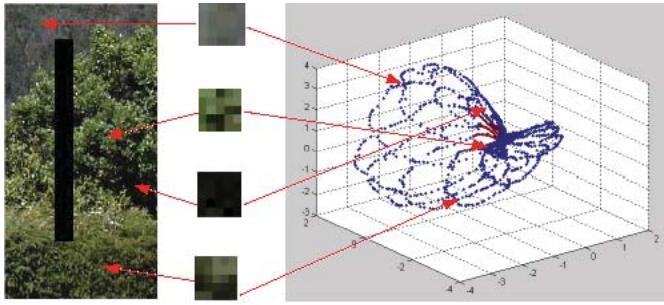


Fig. 3. Image and its embedding. Patches with similar texture are mapped to close locations on the manifold, using an embedding based on LBP texture descriptors [28].

Equation 10 is applied to each eigenvector  $\psi_i$  separately, and while this approach is inferior to using multidimensional nonlinear variational schemes, it emphasizes the simplicity of the inpainting problem, as formulated in the embedding domain. We also applied Criminisi's exemplar based approach [11] in the embedding domain to compare against the isotropic interpolation.

Having computed the extrapolated embedding coordinates  $\{\varphi_i(\mathbf{x}_i, j)\}_1^d$  defined over  $H$ , we aim to relate the embedding coordinates to pixel values in the image domain. Thus, each inpainted pixel  $\mathbf{x}_{i,j} \in H$  is represented by its diffusion coordinates  $\mathbf{y}_{i,j} \in \mathbb{R}^d$ . Unfortunately, the Diffusion embedding is nonlinear and noninvertible. Thus, the mapping  $\Psi^{-1} : \mathbf{y}_{i,j} \mapsto \mathbf{x}_{i,j}$  is nonanalytic and should be approximated numerically. For that we retain for each inpainted point  $\mathbf{y}_{i,j} \in \mathbb{R}^d$  its  $K$  nearest neighbors  $Z_{i,j} = \{\mathbf{y}_{i,j}^1, \dots, \mathbf{y}_{i,j}^K\} \in \mathbb{R}^d$  in terms of diffusion coordinates, that correspond to the set of patches  $\hat{\mathbf{p}}_{i,j} = \{\mathbf{p}_{i,j}^1, \dots, \mathbf{p}_{i,j}^K\}$ .

#### A. Approximate Diffusion Mapping Inversion

In order to solve for the best patch candidates that would allow a globally optimal solution, we aim to assign each image location  $(i, j)$  to one of the potential  $K$  matches  $\hat{\mathbf{p}}_{i,j}$ . The pairwise matching potentials quantify the spatial smoothness between neighboring patches inside of  $H$ , while enforcing consistency with the pixels on the boundary  $\partial H$ .

Given the set of  $\{\mathbf{p}_{i,j}^k\}_1^K$  assignments per inpainted pixel  $\mathbf{x}_{i,j} \in H$ , we formulate a discrete variational approach for choosing the optimal inpainted image values. The discrete formulation, where a single patch  $\mathbf{p}_{i,j}^*$  will be fit to each pixel  $\mathbf{x}_{i,j}$ , allows to overcome the over-smoothing effect when inpainting textured images using linear combinations of image patches. For that we induce smoothness in the image domain by minimizing the discrepancies between overlapping image patches

$$\left\{\mathbf{p}_{i,j}^*\right\} = \arg \min_{\mathbf{p}_{i,j}} \sum_{i,j} \sum_{k_1, k_2=-1}^1 \|\mathbf{p}_{i,j} - \mathbf{p}_{i+k_1, j+k_2}\|^2 \quad (11)$$

$$\text{s.t. } \mathbf{p}_{i,j} \in \left\{\mathbf{p}_{i,j}^k\right\}_1^K$$

Equation 11 implies that for each image location  $(i, j)$  we aim to choose a patch  $\mathbf{p}_{i,j}^*$  out of the  $K$  possible patches

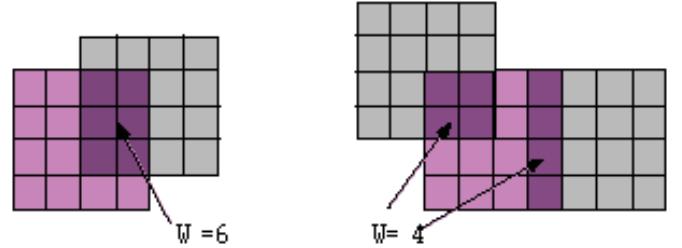


Fig. 4. Similarity between image patches. We compute the SSD over the common support  $\mathbf{p}_{i,j} \cap \mathbf{p}_{i',j'}$ , and normalize by its area  $|\mathbf{p}_{i,j} \cap \mathbf{p}_{i',j'}|$ .

$\hat{\mathbf{p}}_{i,j} = \{\mathbf{p}_{i,j}^1, \dots, \mathbf{p}_{i,j}^K\}$ , such that the overall discrepancy over the hole  $H$  is minimized. This is a pairwise assignment problem that can be formulated as

$$\mathbf{v}^* = \arg \max_{\mathbf{v}} \mathbf{v}^T \mathbf{C} \mathbf{v}, \mathbf{v} \in \{0, 1\}^{|H|K} \quad (12)$$

where  $|H|$  is the number of (unknown) pixels in  $H$ .  $\mathbf{v}$  is a row vectorized replica of the assignment matrix  $\mathbf{V} \in \{0, 1\}^{|H| \times K}$  such that  $v_{i,j} = 1$  implies that the location  $i$  corresponds to the  $j$ 'th patch in  $\hat{\mathbf{p}}_{i,j}$ . The pairwise assignment weight matrix  $\mathbf{C}$  is computed such that

$$c_{((i-1)k+j), ((i'-1)k+j')} = \exp(-\Delta(\mathbf{p}_{i,j} - \mathbf{p}_{i',j'}) / \sigma^2). \quad (13)$$

where  $\Delta(\mathbf{p}_{i,j} - \mathbf{p}_{i',j'})$  is the normalized sum of square differences (SSD) between the common support of the patches shown in Fig. 4.  $\Delta(\mathbf{p}_{i,j} - \mathbf{p}_{i',j'})$  is computed by

$$\Delta(\mathbf{p}_{i,j} - \mathbf{p}_{i',j'}) = \sum_{k_1, k_2 \in \mathbf{p}_{i,j} \cap \mathbf{p}_{i',j'}} \frac{(\mathbf{p}_{i+k_1, j+k_2} - \mathbf{p}_{i'+k_1, j'+k_2})^2}{|\mathbf{p}_{i,j} \cap \mathbf{p}_{i',j'}|} \quad (14)$$

where  $|\mathbf{p}_{i,j} \cap \mathbf{p}_{i',j'}|$  is the area of the common support (depicted in dark purple in fig. 4).

Note that we replaced the minimization in Eq. 11 with the maximization in Eq. 12, which allows to weigh down the numerical effect of outlier assignments that might dominate the formulation in Eq. 11. Given the solution of Eq. 12, we paste the chosen candidates at the corresponding missing pixels in  $H$ . As patches overlap, the overlapping pixels are blended by averaging.

Optimizing Eq. 12 is denoted as the *Quadratic Assignment Problem* (QAP) that is known to be *NP-hard*. Hence, the exact inference of the assignments is intractable and we resort to a suboptimal solution based on spectral relaxation, where we relax the discrete optimization problem in Eq. 12, to an optimization problem with respect to a continuous variable  $\mathbf{s}$

$$\mathbf{s}^* = \arg \max_{\mathbf{s}} \frac{\mathbf{s}^T \mathbf{Cs}}{\mathbf{s}^T \mathbf{s}}, \mathbf{s} \in \mathbb{R}^{|H|k}. \quad (15)$$

This relaxation was used in various computer vision fields [21], [29]–[31], and a probabilistic interpretation was first proposed by Zass and Shashua [32] and then extended by Chertok and Keller [33].

The r.h.s. of Eq. 15 is a Rayleigh quotient, and thus,  $\mathbf{s}^*$  can be computed as the eigenvector corresponding to the leading eigenvalue of  $\mathbf{A}$ . The binary assignment vector  $\mathbf{v}^*$  is computed by applying a discretization procedure to  $\mathbf{s}^*$  maximizing the

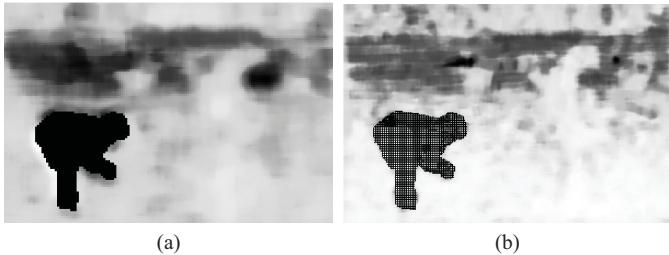


Fig. 5. Diffusion embedding at multiple image resolution scales. (a) Embedding eigenvector of the lower scale image. (b) Embedding eigenvectors of a high-resolution image replica.

sum of the chosen entries in the vector  $\mathbf{s}^*$ . This was shown in [33] to correspond to the maximum likelihood inference of the assignments, under an implicit probabilistic model. As there are no assignment constraints, we apply a greedy discretization scheme, where the relaxed assignment vector  $\mathbf{s} \in \mathbb{R}^{|H|k}$  is reshaped as an assignment matrix  $\mathbf{S} \in \mathbb{R}^{|H| \times k}$ . Thus, each row of  $\mathbf{S}$  corresponds to the assignments of the set of patches  $\hat{p}_{i,j}$  to a particular pixel  $\mathbf{x}_{i,j} \in H$ , and we choose the assignment corresponding to the maximal entry in the corresponding row of  $\mathbf{S}$ .

### B. Multiscale Formulation

We extend the proposed scheme by deriving a multiscale inpainting scheme, using a Gaussian pyramid of images  $I_0, I_1, \dots, I_m$ . The inpainting is applied at the coarsest scale  $I_m$ , and the recovered patches  $h_m$  are used as anchors in the inpainting of  $I_{m-1}$ . The patches in  $I_{m-1}$  whose centers correspond to the locations of  $h_m$  are added to the set  $\partial H_{m-1}$ , that is the set of embedding constraints in scale  $m-1$ . This approach is depicted in Fig. 5, where the lower resolution scale is shown in Fig. 5(a), while the embedding at the higher scale [Fig. 5(b)] utilizes the grid-like anchor embedding points computed in the lower resolution scale.

### C. Affinity Measures

In order to achieve a suitable Diffusion embedding as discussed in Section II-B, one has to choose an appropriate affinity measure to define the Diffusion kernel in Eq. 2. The chosen metric reflects the sort of similarity we are interested in, and the resulting image manifold. For instance, using color descriptors will result in an embedding parameterizing the color manifold of the image. In contrast, in image inpainting we found it most useful to use texture descriptors.

This reflects the relationship between the Diffusion embedding, the heat kernel and the corresponding PDEs. Implying that previous work based on the variational (and PDE-based) formulations, were implicitly operating on an image manifold related to its intensity values. Indeed, such schemes [5] might fail when inpainting textured image regions, and more elaborate schemes that explicitly combine structure and texture inpainting [17] were derived.

We used two texture descriptors: the first being the Local Binary Pattern (LBP) descriptor [28] that characterizes the

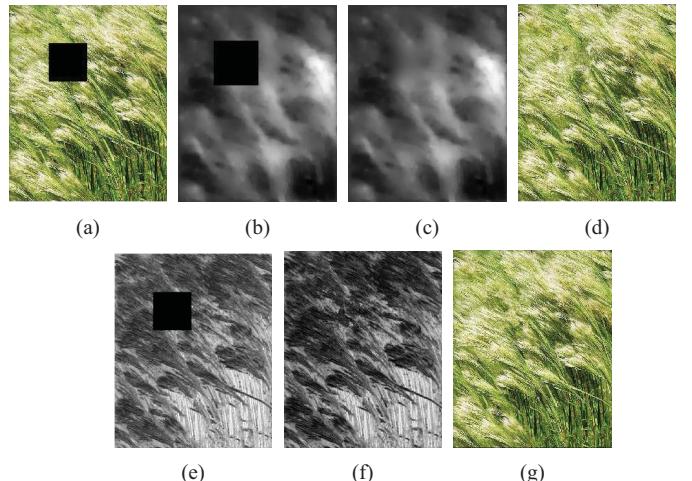


Fig. 6. Comparing the embedding produced using the LBP and Wexler *et al.* [12] texture descriptor. The LBP produces a smoother embedding that can be inpainted using an isotropic PDE, while Wexler's descriptor requires the use of the best-exemplar [11] for the inpainting. (a) Source image. (b) LBP-based leading eigenvector  $\psi_1$ . (c) LBP-based  $\psi_1$  inpainted by isotropic PDE. (d) LBP-based inpainted image. (e)  $\psi_1$  based on the texture descriptor of [12]. (f)  $\psi_1$  based on [12] inpainted by the best-exemplar. (g) Best-exemplar based result.

image texture in the vicinity of each pixel. We use this descriptor to define the similarities between patches in the image. The use of LBP results in a spatially smooth eigenvectors corresponding to the spatial similarity of the image texture. We also used the texture descriptor proposed by Wexler *et al.* [12], consisting of  $(R, G, B, I_x, I_y)$ , being the color channels and the image intensity gradients. This affinity measure results in a less smooth embedding than the one based on LBP, yet it is smoother than the source image.

We compare the embedding and inpainting results of the two descriptors in Fig. 6(b) and (e). The embedding produced by the LBP is smoother than that of Wexler's feature, and while  $\psi_1$  in Fig. 6(b) can be inpainted using isotropic PDE, the eigenvector in 6(e) is not as smooth, and we applied the Exemplar-based approach [11]. Figures 6(d) and (g) show the comparable inpainting results. Our approach is summarized in Algorithm 1.

### D. Discussion

The proposed scheme relates to variational approaches, as it utilizes the induced smoothness of the image embedding. By choosing different affinity measures (and image descriptors) we can formulate the inpainting process with respect to different image manifolds, such as texture, color, and edges. Previous variational schemes were implicitly restricted to using the intensity or color manifold of the image. Thus, our approach allows to apply variational schemes on generalized manifolds such as textured images. In a broader view, the proposed scheme can be applied to the interpolation of generalized signals and datasets, by choosing appropriate affinity measures. For instance, it can be applied to inpainting/interpolating tabular data, where contemporary variational schemes are inapplicable. Initial results of this approach were presented by Liberti *et al.* [34].

**Algorithm 1** Image Completion

---

```

1: Input: image  $I$ , hole mask  $H \subset I$ , manifold learning set  $\bar{H}$ .
2: Compute an image feature for each pixel in  $\bar{H}$ .
3: Compute the affinity matrix  $\mathbf{W}$ , where  $w_{ij} = \exp(-d_{ij}^2/\sigma^2)$ .
4: Compute  $\Psi(\bar{H}) = \{\psi_i(\mathbf{x}_{i,j})\}_1^d$  leading eigenvectors of  $\mathbf{W}$ .
5: Inpaint embeddings  $\Psi(H)$  within the hole  $H$ .
6: for all  $x_{ij} \in H$  do
7:   Search for K-nn points in  $\Psi(\bar{H})$ :  $Z_{i,j} = \{\mathbf{y}_{i,j}^1, \dots, \mathbf{y}_{i,j}^K\} \in \mathbb{R}^d$ , and corresponding patches
       $\hat{p}_{i,j} = \{\mathbf{p}_{i,j}^1, \dots, \mathbf{p}_{i,j}^K\}$ .
8: end for
9: Solve the inverse Diffusion mapping using spectral relaxation.
10: Compose the result image by blending the chosen patches.

```

---

In contrast, exemplar based schemes forgo the smoothness property of the manifold, and even the assumption that a manifold exists, by inpainting using nearest neighbor sampling. Thus, such schemes are able to inpaint textured (“high dimensional”) image regions and tabular data. The downside is that they can only be applied locally, causing the loss of the global optimality of the inpainting. This is exemplified by the propagation of errors in early exemplar based schemes, that was solved by the detection of large scale image structures [11] and iterative approaches [12].

The assumption of manifolds smoothness and low dimensionality is also related to the sparse representations schemes that were applied to image inpainting [16], [27], where the image was represented by a set of patches, approximated by a sparse set of coefficients and a corresponding sparse dictionary. It is interesting to note that both approaches utilize the notion of low-dimensionality manifested by the small number of coefficients and dictionary atoms/eigenvectors used to represent the data. They differ on their approach for deriving the low dimensional representation. While the sparse representation approach explicitly aims for the sparse representation using  $L_1$  minimization, the Diffusion approach assumes the data is low dimensional and recovers the corresponding low dimensional parametrization.

Complexity-wise, our scheme requires a single computation of the distances between the patches in the learning set  $\bar{H}$  to compute the Diffusion embedding. The upside is that this computation is conducted as a preprocessing step, while in exemplar based schemes such as the Best-exemplar and Wexler’s work, the distances between the active set of patches and the learning set are computed in each iteration.

#### IV. EXPERIMENTAL RESULTS

In this section we detail the experimental results exemplifying the proposed inpainting scheme. We applied our schemes to real images used in contemporary work. In Section IV-A we show the diffusion domain inpainting results using an isotropic heat equation (IHE), while the results of applying the best-exemplar in the diffusion domain are shown in Section IV-B.

The robustness to noise of the proposed scheme is verified and analyzed in Section IV-C, and we study the effectiveness of common state-of-the-art image quality measures in quantifying the quality of image inpainting results in Section IV-D. Failure cases are presented and discussed in Section IV-E, and we detail the criteria for setting the algorithm’s parameters, and other implementation issues, as well as the timing results in Section IV-F.

In all examples, the proposed scheme was implemented using a Gaussian kernel, and  $K = 10$  candidate patches used as input to the combinatorial solver. We compared against state-of-the-art inpainting schemes, which include Criminisi *et al.*’s Best exemplar<sup>1</sup> [11], the PDE based approach of Tschumperle [22]<sup>2</sup> (using 200 iterations and the default values of the other parameters) and Darren Lafreniere’s implementation<sup>3</sup> of Komodakis *et al.*’s approach [26] (using the “full” mode that provides improved inpainting quality). Where possible, we applied these approaches and ours to the images used by Wexler *et al* [12], and thus compared against their reported results. Last, we inpainted the images using the IHE in the image domain. This emphasizes the benefits of operating in the Diffusion domain, that is the focal point of our work. The inpainting quality assessments made in this section are in essence subjective, but we did our best to be as impartial as possible.

##### A. Isotropic Heat Equation Inpainting in the Diffusion Domain

In this section we report the results of applying our approach using IHE inpainting in the diffusion domain. For that we implemented an isotropic heat equation solver that was applied to inpaint the embedding vectors which are missing the same domain  $H$  as the image.

Figs. 7 and 8 shows two examples of inpainting using the LBP as a texture feature, where the features were computed using  $15 \times 15$  histograms of the pointwise LBP measure [28]. Although the source images (Fig. 8 in particular) are textured, the embedding vectors [Figs. 7(b) and 8(b)] are smooth, and are well inpainted by a simple IHE scheme. In contrast, when applying the IHE and the PDE-based scheme of Tschumperle [8] in the image domain, both struggle to inpaint the two images, resulting in over-smoothing in Figs. 7(d) and 8(d). Fig. 8(d) and (g) show that PDE based schemes are able to inpaint narrow objects such as cracks or cables, where the over-smoothing is less evident. The proposed scheme seems to provide the best inpainting results in Fig. 7, while being as good as the Best exemplar in Fig. 8.

##### B. Diffusion Domain Inpainting Using the Best-Exemplar

In this section we apply Criminisi’s Best exemplar inpainting approach in the Diffusion domain, and show it provides improved results compared to its (common) use in the spatial domain. Figures 9–11 present inpainting results, where we used Wexler’s five-dimensional image feature [12], instead of the LBP, and the best-exemplar rather than the IHE to

<sup>1</sup> Available at: <http://www.cc.gatech.edu/~sooraj/inpainting/>

<sup>2</sup> Available at: <http://gmic.sourceforge.net/>

<sup>3</sup> Available at: <http://www.lafarren.com/image-completer/>

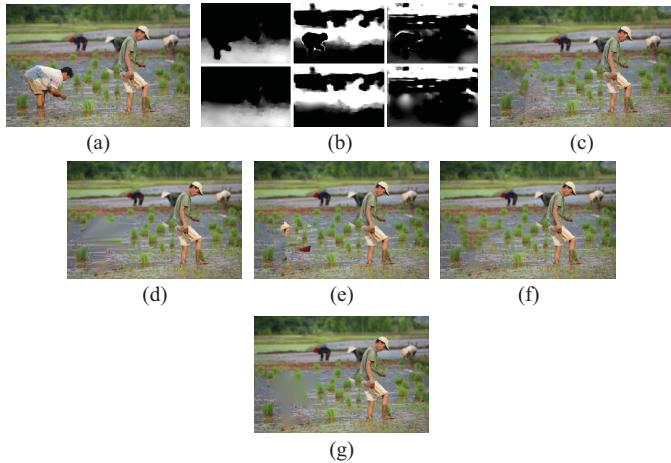


Fig. 7. Inpainting results of the *Rice field* image. The embedding was computed using the LBP texture descriptor. (a) Original image. (b) Leading eigenvectors and their inpaintings. Upper row: the leading embedding eigenvectors with the unknown part  $H$ . Bottom row: the inpainted eigenvectors. (c) Proposed scheme. (d) Tschumperle *et al.* [22]. (e) Best-exemplar [11]. (f) Komodakis *et al.* [26]. (g) Isotropic heat equation.

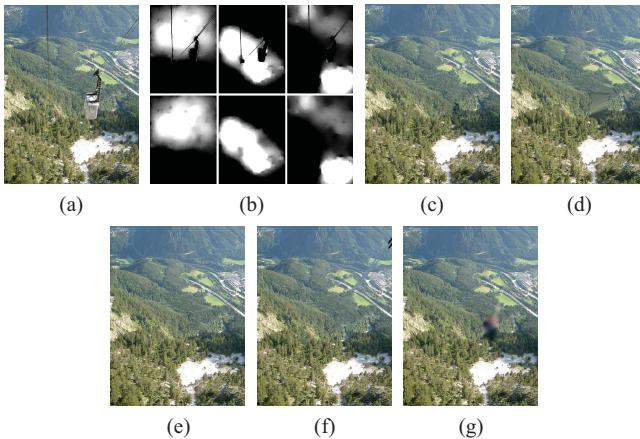


Fig. 8. Inpainting results of the *Cable* image. The embedding was computed through the LBP texture descriptor. (a) Original. (b) Leading eigenvectors and their inpaintings. Upper row: the eigenvectors with the unknown part  $H$ . Bottom row: the inpainted eigenvectors. (c) Proposed scheme. (d) Tschumperle *et al.* [22]. (e) Best-exemplar [11]. (f) Komodakis *et al.* [26]. (g) Isotropic heat equation.

inpaint the eigenvectors. In Fig. 9 the proposed scheme proved superior to the other approaches, while in Figs. 10 and 11 our results are on par with those of the Best exemplar and Komodakis *et al.*. These examples emphasize the attributes of the different schemes. In Fig. 9 the unknown region has a global structure that is best inpainted by a globally optimal scheme such as ours. In contrast, the missing region in Figs. 10 and 11 is made of isotropic texture that does not require global optimization, and is thus well inpainted by all schemes. The embedding smoothness is particularly evident in Fig. 11.

In Fig. 12 we show another example of an image consisting of texture with global structures, that are the dark lines. As in Fig. 9, The proposed schemes provides the best inpainting results as it is able to combine texture synthesis with global optimization. In Fig. 12(b) we show the three leading

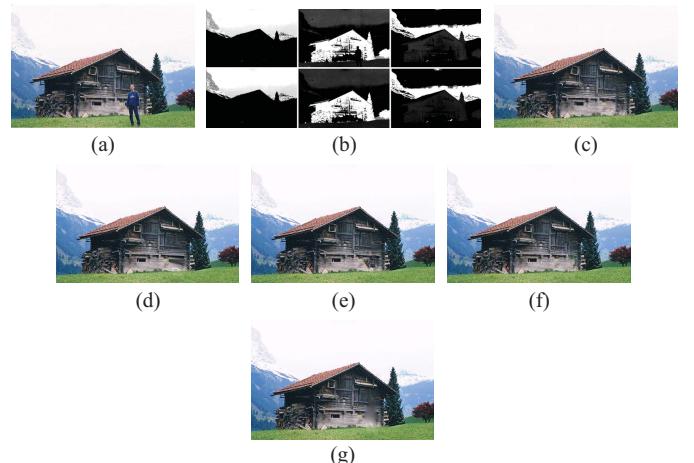


Fig. 9. Image completion through the 5-D texture descriptor of [12] and best-exemplar inpainting of the eigenvectors. (a) Original image. (b) Leading embedding eigenvectors and their inpaintings. (c) Proposed scheme. (d) Tschumperle *et al.* [22]. (e) Best exemplar [11]. (f) Komodakis *et al.* [26]. (g) Isotropic heat equation.

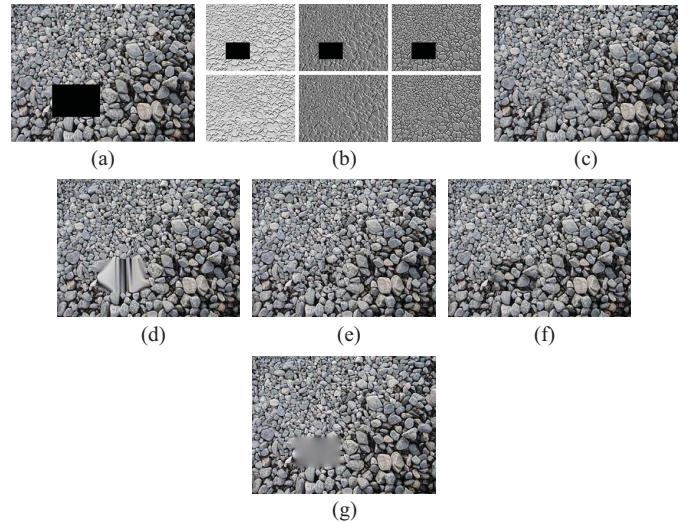


Fig. 10. Image completion using the 5-D texture descriptor of [12] and best-exemplar inpainting of the eigenvectors. (a) Original image. (b) Leading embedding eigenvectors and their inpaintings. (c) Proposed scheme. (d) Tschumperle *et al.* [22]. (e) Best exemplar [11]. (f) Komodakis *et al.* [26]. (g) Isotropic heat equation. For each image, we report the image quality measures discussed in Section IV-D.

embedding eigenvectors, that are the first three Diffusion coordinates of each pixel. It follows, that by using texture features, the resulting embedding is smooth, making the inpainting easier. Indeed, applying the Best-exemplar in the embedding domain yields improved results compared the applying it in the spatial domain. In Fig. 12(c) we depict the candidate patches and their locations in the image.

We were unable to obtain an implementation of Wexler *et al.*'s scheme [12] to compare against. Hence, we inpainted the *Bungee* image used in [12] and report the results in Fig. 13. We used the five-dimensional texture features used in [12] and inpainted the embeddings using the Best-exemplar. The eigenvectors in Fig. 13(b) are relatively smooth and are well inpainted. We achieve comparable results to [12], while

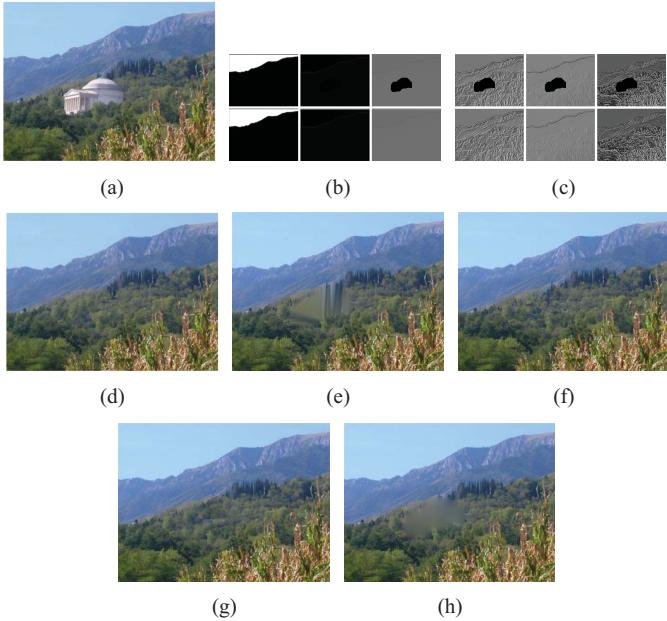


Fig. 11. Image completion using the 5-D texture descriptor of [12] and best-exemplar inpainting of the eigenvectors. (a) Original image. (b) Eigenvectors  $\psi_1/\psi_3$  and their inpaintings. (c) Eigenvectors  $\psi_{10}/\psi_{12}$  and their inpaintings. (d) Proposed scheme. (e) Tschumperle *et al.* [22]. (f) Best-exemplar [11]. (g) Komodakis *et al.* [26]. (h) Isotropic heat equation.

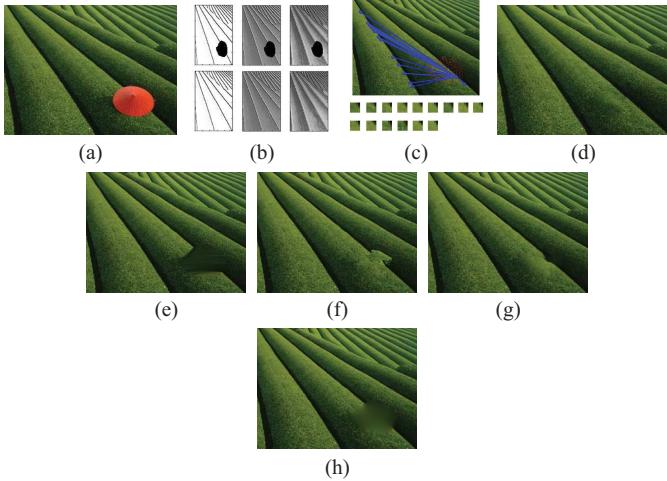


Fig. 12. Image completion using the 5-D texture descriptor of [12] and best-exemplar [11] inpainting of the eigenvectors. (a) Original image. (b) Leading embedding eigenvectors and their inpaintings using the best-exemplar approach. (c) Patch candidates. Red dots: central pixels of the hole patches. The location of the  $K = 15$  candidates patches are shown. (d) Proposed scheme. (e) Tschumperle *et al.* [22]. (f) best-exemplar [11]. (g) Komodakis *et al.* [26]. (h) Isotropic heat equation.

being superior to the other schemes. Similar to ours, Wexler's algorithm aims for global optimality, by iterating the image inpainting as detailed in Section II-A, while our algorithm runs a single iteration.

### C. Robustness to Noise

The robustness of the proposed scheme to additive image noise, was tested by adding White Gaussian Noise (WGN) to the inpainted image. The results are depicted in Fig. 14, where we show the inpainting results for varying WGN levels

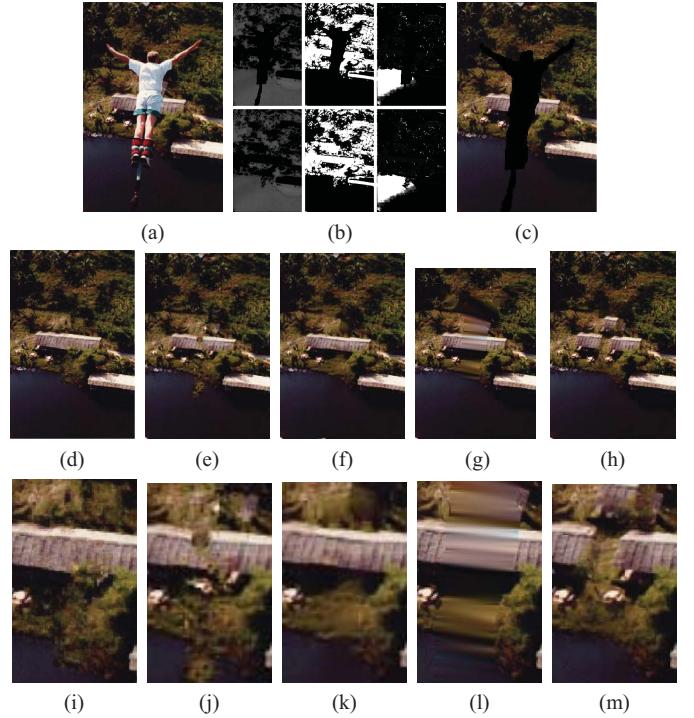


Fig. 13. Inpainting results of the *Bungee* image. (a) Original image. (b) Leading embedding eigenvectors and their inpaintings using a 5-D texture feature [12], and Criminisi's best-exemplar inpainting of the embedding vectors. (c) Masked image. (d) Proposed scheme. (e) Best-exemplar result [11]. (f) Wexler *et al.* [12] result. (g) Tschumperle *et al.* [22]. (h) Komodakis *et al.* [26]. (i)–(m) Zoom into the inpainted area of (d–h), respectively.

$\sigma_N = \{0.1, 0.2, 0.3, 0.4\}$ . As the images are inpainted using the noisy patches, the resulting images are also noisy. But the inpainting is numerically stable.

This robustness can be attributed to two underlying factors. First, the patch affinity measure discussed in Section III-A and illustrated in Fig. 4, averages the intensity differences over the common patch, attenuating the noise effect. Second, the robustness of spectral embedding schemes is well studied, due to their proliferation (Isomap, LLE, Hessian-LLE, Laplacian Eigenmaps, Diffusion Maps, to name a few), and successful use in a gamut of applications. The noisy distance matrix to be embedded is a low-rank matrix, whose eigendecomposition is known to be robust, due to Wigner's Semicircle Law [35], which implies that adding noise to p.s.d. matrices results in a semi-circle shaped noise spectrum. As the noise increases, the semi-circle noise spectrum moves towards the leading eigenvalues, and when the noise energy reaches a certain threshold, the semi-circle crosses over into the leading eigenvalues [36]. The spectrum of affinity matrices were studied by El-Karoui [37] who showed that smooth kernels (such as the Gaussian kernel) lead to the Marcenko-Pastur distribution, and can be analyzed via Wigner's Semicircle Law. Thus, below a certain level of noise energy, the computation of embedding vectors  $\{\psi_i(\mathbf{x}_{i,j})\}_1^d$  is invariant to the noise. As the noise energy increases, we encounter a crossover phenomenon, where the eigenvectors related to the noise create a linear mixture with the embedding vectors related to the underlying data, thus corrupting them.

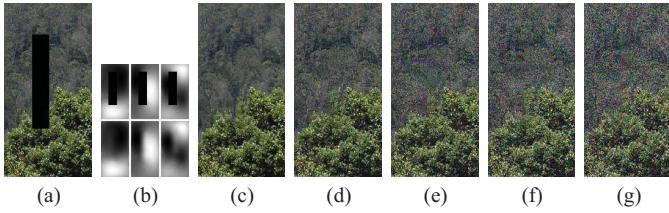


Fig. 14. Inpainting results for noisy images. The leading embedding eigenvectors in (b) relate to  $\sigma_N = 0.4$ . (a) Source. (b) Embedding. (c)  $\sigma_N = 0$ . (d)  $\sigma_N = 0.1$ . (e)  $\sigma_N = 0.2$ . (f)  $\sigma_N = 0.3$ . (g)  $\sigma_N = 0.4$ .

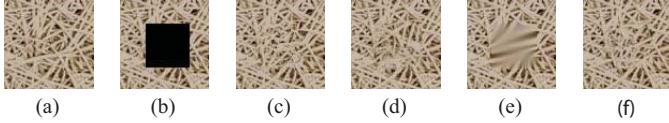


Fig. 15. SSIM, VIF, and DIIIVINE image quality scores. The quality range of the SSIM and VIF is  $[0, 1] \sim [\text{worst}, \text{best}]$ , but for the DIIIVINE measure, the lower the score the better. (a) Original images. (b) Missing parts to be inpainted. (c) Best-exemplar result [11]. (d) Komodakis et al. [26]. (e) Tschumperle et al. [22]. (f) Proposed scheme.

#### D. Image Inpainting Quality Measures

The quantification of image quality inpainting is a fundamental issue that is often overlooked, as it is common to exemplify the inpainting quality by visual comparisons. We applied three state-of-the-art image quality measures, Structural Similarity (SSIM) [38], Visual Information Fidelity (VIF) [39] and Distortion Identification-based Image Verity and Integrity Evaluation (DIIIVINE) [40], whose code is publicly available, to some of the inpainted images. The results are reported in Figs. 10 and 15, where for both SSIM and VIF, the quality range  $[0, 1] \sim [\text{worst}, \text{best}]$ , while for the DIIIVINE measure, the score range is  $[0, 100] \sim [\text{best}, \text{worst}]$ .

Unfortunately, it seems that these image quality measures are inconsistent when applied to the inpainting results. Thus, in Fig. 10, the SSIM and VIF scores seem uncorrelated to the visual output. Higher SSIM and VIF scores were given to Fig. 10(d) than to Figs. 10(c), 10(e), and 10(f) although these seem (to us) to be the more visually pleasing. The same applies to the DIIIVINE measure that seems to favor Fig. 10(d). The same can be said of the results in Fig. 15 where Fig. 15(e) leads the SSIM and VIF scores, while being the worse on the DIIIVINE scale.

#### E. Failure Cases

The scheme we propose might fail under certain conditions. We present two such examples in Fig. 16. The first is the *Parrot* image in Figs. 16(a)–16(d). The failure is due to the widening of the unknown regions  $H$  in the upper row of Fig. 16. The widening occurs due to the use of  $15 \times 15$  image patches to define the LBP descriptors. Thus, in order to compute the LBP descriptor on the boundary  $\partial H$ , we had to widen  $H$  by 30 pixels overall. In contrast, PDE-based schemes use a two-pixels wide boundary  $\partial H$ , but might fail to inpaint textured areas. This is depicted in Fig. 16(d), where we applied the PDE-based approach of Tschumperle et al. [22]. Although the inpainting shows little artifacts, close visual inspection reveals that the inpainted segment is over-smoothed and that

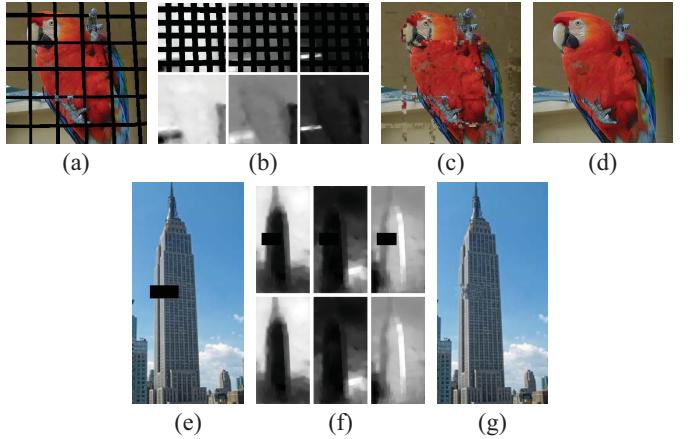


Fig. 16. Inpainting failure cases. (a) *Parrot* image with fence removed. (b) Leading embedding eigenvectors and their inpaintings. (c) Inpainting results by the proposed scheme. (d) Tschumperle et al. [22]. (e) *Building* image with a missing block. (f) Leading embedding eigenvectors and their inpaintings. (g) Inpainting results by the proposed scheme [22].

texture fine details are lost. A similar phenomenon occurs in the inpainting of the building shown in Fig. 16(e)–(g), where the isotropic PDE-based approach fails to preserve the large scale geometric structure. This is evident in the form of the leading embedding eigenvector shown in Fig. 16(f).

#### F. Parameters Settings and Implementation Issues

The proposed inpainting schemes whose results are shown in Sections IV-A and IV-B, perform differently when applied to different image types. The variational approach we applied is based on an isotropic heat equation, and is thus suitable for textured images not having large scale geometrical structures across the hole. For instance, the IHE approach is inappropriate for the eigenvectors image shown in Fig. 12(b). The inpainted texture might be heterogeneous (e.g. water, grass, and mud in Fig. 7), and still produce a visually plausible image. We utilized it alongside the LBP texture descriptor, that is best suitable for textured image regions.

The exemplar-based approach preserves large scale geometric structures in the eigenvector images. We found the LBP to be inefficient in such scenarios, as it encodes local histograms of texture data, and thus obscures the geometric structure. Wexler's descriptor matches natural patches emphasizing high gradients, allowing to preserve gradient and geometrical structures.

Our approach utilizes a few parameters that we found to be stable over the inpainted image set. The dimensionality of the embedded space  $d$ , is related to the intrinsic dimensionality of the image. For heterogenous images consisting of different textures we set  $d = 15$ . Thus, in Fig. 11(b), the first three eigenvectors characterizing the sky, their complement, and the skyline. For  $d = 3$  any patch from the “nonsky” part of the image could be chosen to fill the hole. More detailed parts can be seen in Fig. 11(c), as higher order eigenvectors may contain mostly noise.

The Gaussian kernel width  $\sigma$  used by the diffusion maps, defines the nearest neighbor structure of the graph, and

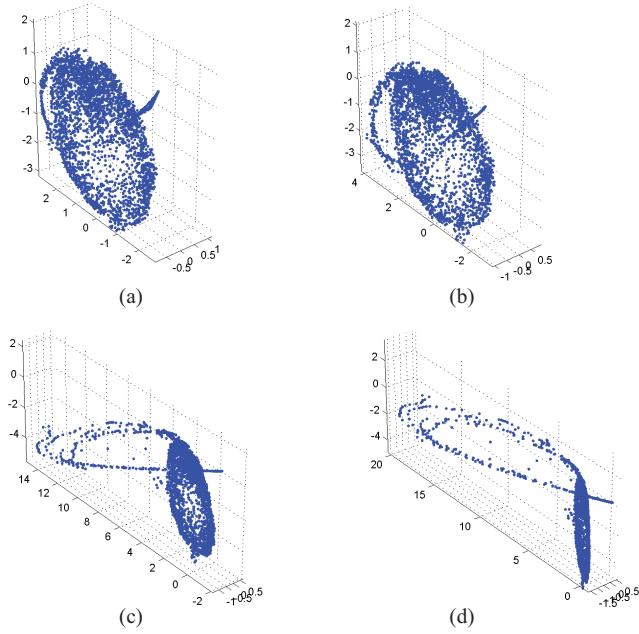


Fig. 17. Embedding of the patches of the *Bungee* image for different kernel bandwidths  $\sigma$ . (a)  $\sigma = d_{\max}$ . (b)  $\sigma = d_{\max}/3$ . (c)  $\sigma = d_{\max}/6$ . (d)  $\sigma = d_{\max}/8$ .

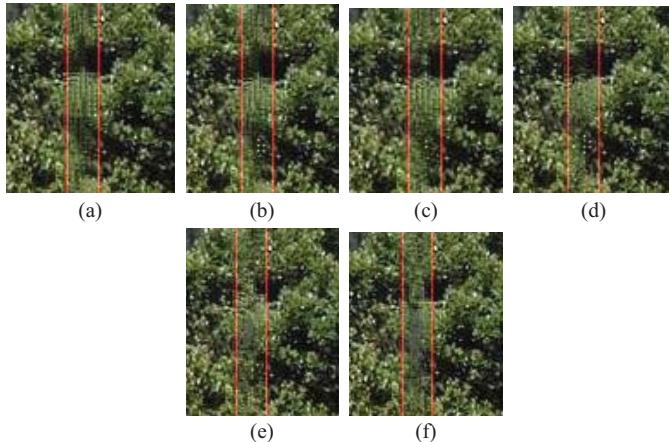


Fig. 18. Inpainting result of the proposed scheme with respect to different  $K$  nearest neighbors in the assignment problem (Section III-A). (a)  $k = 1$ . (b)  $K = 2$ . (c)  $K = 3$ . (d)  $K = 5$ . (e)  $K = 10$ . (f)  $K = 20$ .

therefore the topology of the manifold of the diffusion map. While very small values of  $\sigma$  may disconnect branches of the graph, large values may connect dissimilar patches. In practice, the proposed scheme is robust to the choice of  $\sigma$ , as there is a wide range of its values that yields similar, close to optimal results. We used  $\sigma = d_{\max}/3$  where  $d_{\max}$  is the maximal patch distance within the learning set. Similar results can be achieved by using the median as well. In Fig. 17 we show the embeddings of the *Bungee* image using varying  $\sigma$  values. These may differ in shape but have similar global structure. The K-nn search in the embedding domain yields similar results for different values of  $\sigma$ , as long as  $\sigma$  is not small enough to break the graph topology.

We exemplify the effect of varying the number of candidate patches in Fig. 18 for  $K = \{1, 2, 3, 5, 10, 20\}$ . Setting  $K = 1$

[Fig. 18(a)] implies that the most similar patch in terms of Diffusion distance is used, but no spatial image smoothness is induced. As  $K$  increases, the quality of the inpainting improves, as depicted in Figs. 18(b)–(e). In assignment problems where the number of possible assignments  $K \geq 10$ , the performance of the spectral approach (Section III-A) deteriorates, thus reducing the inpainting quality as shown in Fig. 18(f) for  $K = 20$ .

The proposed scheme was implemented in MATLAB and the timing was measured on a Quad4-based computer running at 2.6 GHz. When inpainting a  $200 \times 200$  image, the computation of the sparse distance matrix between the LBP features, used to compute the embedding required 11 minutes. Applying the IHE on the Diffusion embedding vectors took a minute, while running the Best-exemplar inpainting required three minutes. The approximate inverse mapping using the spectral relaxation lasted three minutes on average. Thus, the overall inpainting lasts 15 and 17 minutes, for the IHE and Best Exemplar-based schemes, respectively. This timing can be significantly shortened by improving the computation of the patch distance matrix using the Patchmatch approach [3] and its variations.

## V. CONCLUSION

In this work we proposed a Diffusion based framework for image completion. The crux of our approach is to utilize the Diffusion embedding to induce application specific smoothness over the inpainted image. The induced smoothness is manifested by the smoothness of the embedding eigenvectors, when computed using appropriate affinity measures, such as the LBP texture features. Thus, we can apply PDE-based inpainting schemes that are commonly only applicable to smooth images, to textured images. In particular, such schemes allow to achieve global inpainting optimality and can be applied to inpainting large-scale holes. We also showed that the use of the Best-exemplar in the diffusion domain, rather than in image intensity domain, yields improved inpainting results. In order to compute an approximate inverse-diffusion mapping, we introduce a novel approach based on discrete optimization, with a corresponding spectral relaxation. The proposed inpainting scheme compares favorably with previous state-of-the-art methods.

## REFERENCES

- [1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, “Image inpainting,” in *Proc. 27th Annu. SIGGRAPH Conf., Comput. Graph. Interact. Tech.*, 2000, pp. 417–424.
- [2] S. Avidan and A. Shamir, “Seam carving for content-aware image resizing,” *ACM Trans. Graph.*, vol. 26, no. 3, pp. 1–9, Jul. 2007.
- [3] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman, “Patchmatch: A randomized correspondence algorithm for structural image editing,” *ACM Trans. Graph.*, vol. 28, pp. 24:1–24:11, Jul. 2009.
- [4] P. Pérez, M. Gangnet, and A. Blake, “Poisson image editing,” in *Proc. ACM SIGGRAPH Papers*, 2003, pp. 313–318.
- [5] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, “Image inpainting,” in *Proc. 27th Annu. Conf. Comput. Graph. Int. Tech.*, 2000, pp. 417–424.
- [6] M. Bertalmio, “Strong-continuation, contrast-invariant inpainting with a third-order optimal PDE,” *IEEE Trans. Image Process.*, vol. 15, no. 7, pp. 1934–1938, Jul. 2006.
- [7] J. Shen, S. H. Kang, and T. F. Chan, “Euler’s elastica and curvature-based inpainting,” *SIAM J. Appl. Math.*, vol. 63, no. 2, pp. 564–592, 2003.

- [8] D. Tschumperlé, "Fast anisotropic smoothing of multi-valued images using curvature-preserving pde's," *Int. J. Comput. Vis.*, vol. 68, no. 1, pp. 65–82, 2006.
- [9] J. Shen, "Inpainting and the fundamental problem of image processing," *SIAM News*, vol. 36, no. 5, pp. 1–4, Jun. 2003.
- [10] A. Efros and T. Leung, "Texture synthesis by non-parametric sampling," in *Proc. IEEE Int. Conf. Comput. Vis. Proc. 7th*, vol. 2, Mar. 1999, pp. 1033–1038.
- [11] A. Criminisi, P. Pérez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1200–1212, Sep. 2004.
- [12] Y. Wexler, E. Shechtman, and M. Irani, "Space-time completion of video," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 3, pp. 463–476, Mar. 2007.
- [13] I. Dror, D. Cohen-Or, and H. Yeshurun, "Fragment-based image completion," *ACM Trans. Graph.*, vol. 22, pp. 303–312, Jul. 2003.
- [14] H. Igehy and L. Pereira, "Image replacement through texture synthesis," in *Proc. Int. Conf. Image Process.*, vol. 3, pp. 186–189, Oct. 1997.
- [15] J. Hays and A. A. Efros, "Scene completion using millions of photographs," *ACM Trans. Graph.*, vol. 26, no. 3, p. 1, Jul. 2007.
- [16] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. Image Process.*, vol. 17, no. 1, pp. 53–69, Jan. 2008.
- [17] A. Bugeau, M. Bertalmio, V. Caselles, and G. Sapiro, "A comprehensive framework for image inpainting," *IEEE Trans. Image Process.*, vol. 19, no. 10, pp. 2634–2645, Oct. 2010.
- [18] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher, "Simultaneous structure and texture image inpainting," *IEEE Trans. Image Process.*, vol. 12, no. 8, pp. 882–889, Aug. 2003.
- [19] R. Coifman and S. Lafon, "Diffusion maps," *Appl. Comput. Harmon. Anal. Spec. Issue Diffus. Maps Wavelets*, vol. 22, no. 1, pp. 5–30, Jul. 2006.
- [20] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Comput.*, vol. 6, pp. 1373–1396, Jun. 2003.
- [21] M. Leordeanu and M. Hebert, "A spectral technique for correspondence problems using pairwise constraints," in *Proc. Int. Conf. Comput. Vis.*, Oct. 2005, pp. 1482–1489.
- [22] D. Tschumperlé and R. Deriche, "Vector-valued image regularization with PDEs: A common framework for different applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 4, pp. 506–517, Apr. 2005.
- [23] S. Masnou and J.-M. Morel, "Level lines based disocclusion," in *Proc. Int. Conf. Image Process.*, Oct. 1998, pp. 259–263.
- [24] O. Le Meur, J. Gautier, and C. Guillemot, "Exemplar-based inpainting based on local geometry," in *Proc. 18th IEEE Int. Conf. Image Process.*, Sep. 2011, pp. 3401–3404.
- [25] S. Xue and Q. Dai, "Nonlinear poisson image completion using color manifold," in *Proc. IEEE Int. Conf. Image Process.*, vol. 4, Oct. 2007, pp. 237–240.
- [26] N. Komodakis and G. Tziritas, "Image completion using efficient belief propagation via priority scheduling and dynamic pruning," *IEEE Trans. Image Process.*, vol. 16, no. 11, pp. 2649–2661, Nov. 2007.
- [27] M. Elad, J.-L. Starck, P. Querre, and D. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," *Appl. Comput. Harmon. Anal.*, vol. 19, no. 3, pp. 340–358, 2005.
- [28] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.
- [29] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.
- [30] A. Egozi, Y. Keller, and H. Guterman, "Improving shape retrieval by spectral matching and meta similarity," *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1319–1327, May 2010.
- [31] M. Chertok and Y. Keller, "Spectral symmetry analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 7, pp. 1227–1238, Jul. 2010.
- [32] R. Zass and A. Shashua, "Probabilistic graph and hypergraph matching," in *Proc. IEEE Comput. Vis. Pattern Recognit. Conf.*, Jun. 2008, pp. 1–8.
- [33] M. Chertok and Y. Keller, "Efficient high order matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2205–2215, Dec. 2010.
- [34] E. Liberty, M. Almagor, S. Zucker, Y. Keller, and R. Coifman, "Scoring psychological questionnaires using geometric harmonics," in *Proc. Snowbird Learn. Workshop*, 2007, pp. 1–2.
- [35] F. Chung, *Spectral graph theory*. AMS, Providence, RI, no. 92, May 1997.
- [36] B. Nadler, "Finite sample approximation results for principal component analysis: A matrix perturbation approach," *Anna. Statist.*, vol. 36, no. 6, pp. 2791–2817, 2008.
- [37] N. El Karoui, "The spectrum of kernel random matrices," *Anna. Statist.*, vol. 38, no. 1, pp. 1–50, 2010.
- [38] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Perceptual image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [39] H. Sheikh and A. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.
- [40] A. Moorthy and A. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3350–3364, Dec. 2011.



**Shai Gepshtein** received the B.Sc. and M.Sc. degrees in electrical engineering from Tel-Aviv University, Tel-Aviv, Israel, in 1998 and 2004, respectively. He is currently pursuing the Ph.D. degree in electrical engineering with the Faculty of Engineering, Bar-Ilan University, Ramat-Gan, Israel.

He was an R&D Engineer from 1998 to 2012.



**Yosi Keller** received the B.Sc. degree in electrical engineering from the Technion-Israel Institute of Technology, Haifa, Israel, in 1994, and the M.Sc. and Ph.D. degrees in electrical engineering from Tel-Aviv University, Tel-Aviv, Israel, in 1998 and 2003, respectively.

He was a Gibbs Assistant Professor with the Department of Mathematics, Yale University, New Haven, CT, from 2003 to 2006. He is an Associate Professor with the Faculty of Engineering, Bar Ilan University, Ramat-Gan, Israel.