

Lost but Found? Harnessing the Internet for Photometric Completion

Pratyush Sahay
Department of Electrical Engineering, IIT Madras, Chennai, India.
pratyush.a.sahay@gmail.com

A.N. Rajagopalan
Department of Electrical Engineering, IIT Madras, Chennai, India.
raju@ee.iitm.ac.in

Abstract

While it is important to digitize heritage sites ‘as is’, building 3D models of damaged archaeological structures can be visually unpleasant due to the presence of large missing regions. This work addresses intensity filling-in, or intensity inpainting, of such large damaged regions post geometric reconstruction. Assuming a Lambertian image formation model, we first establish that patches corresponding to arbitrarily oriented planar regions found in internet images of several archaeological structures with possibly different albedo and observed under varied and uncontrolled illumination lie in a low-dimensional subspace. These are then used for inpainting by modeling the missing region as gross pixel corruptions. The performance of the proposed method along with comparisons are shown on synthetic as well as real data.

1. Introduction

The importance of digitization of heritage, present in the form of archaeological sites, museums showcasing historical artifacts and famous paintings, etc. is well-documented in UNESCO’s draft charter on the preservation of digital heritage [1]. Almost in parallel, a large-scale interest in heritage digitization arose in the vision, graphics, virtual reality and related research communities with the advent of efforts such as The Digital Michelangelo Project [14] and Google Art Project. In addition to preservation, such projects aim to provide a capability to perform a virtual walk-through, enabling internet-based access of rich ‘common’ heritage from across the world.

Due to numerous forces of degradation (natural and man-made) at work, several of these archaeological structures tend to have large damaged or broken regions. The above mentioned projects, in the current state, would show such a naturally existing large missing region ‘as is’. Thus, a natural addition to such heritage visualization applications would be the ability to geometrically and photometrically

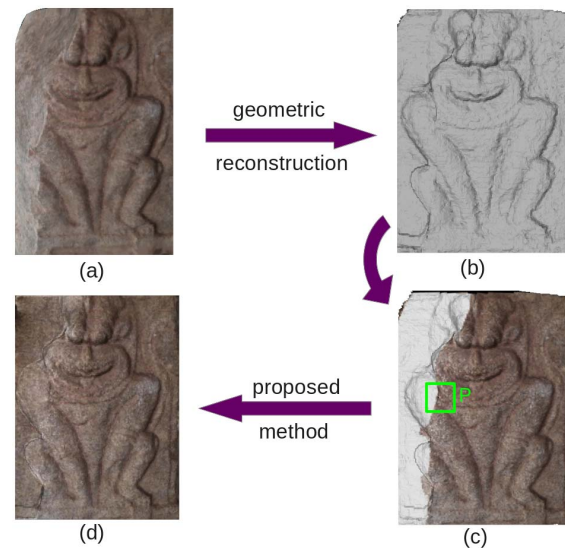


Figure 1. The geometric and photometric hole reconstruction pipeline. (a) Original 3D model, (b) geometrically reconstructed model using [20], (c) patch showing both known and unknown intensity regions, and (d) intensity inpainted 3D model.

reconstruct large missing regions in the rendered 3D models (Fig. 1). Geometric reconstruction of such damaged regions from self-similar examples has been proposed in [20]. However, for a visually pleasing reconstruction and seamless incorporation into such 3D visualization applications, texture filling-in for the reconstructed region is an important module in the missing-region-reconstruction pipeline. Due to the possibility of wide texture variations, such a photometric completion task is by no means trivial.

While [11] attempts to solve a somewhat related problem of shape and texture reconstruction using examples, the restrictive requirements of custom background, illumination and a reference object during image acquisition are difficult to mimic for archaeological data. In contrast, the work described in this paper does away with any such restrictions.

Following the nomenclature introduced in [10], this paper deals with ‘intensity hallucination’ of the geometri-

cally hole-filled regions in damaged 3D models. In our approach, we first establish that a corpus of image patches corresponding to arbitrarily oriented planar regions found locally in *several* Lambertian surfaces, with varied albedo and illumination conditions, exhibit a low-rank structure. A reliable set of such exemplars is easily created using local patches from internet images of popular archaeological structures. These tend to have a large number of publicly shared images taken by people from across the globe, with text searches like ‘archaeology’, ‘archaeology travel’ and ‘archaeology close up’ returning approximately 38527 and 395000 results on Flickr and Google Image Search, respectively. Modeling the missing region in a patch P (Fig. 1 (c)) as gross pixel corruptions, we show that it is possible to simultaneously recover this low-rank subspace and the sparse error matrix using advanced convex optimization techniques introduced in [23], [27]. The recovered low-dimensional subspace inherently performs filling-in of the missing region in patch P . Recovering low-rank matrices, corrupted by gross pixel errors, has been proposed in [17], [23], [27] in which applications such as recovering background from a scene with dynamic foreground, and removing specularities from face images were dealt with. These works make use of the inherent spatial correlation in multiple images of the same object or scene to justify the use of the low-rank matrix model.

The main contributions of this work are:

- showing that image patches corresponding to arbitrarily oriented planar regions found locally in *several* Lambertian surfaces, with varied albedo and lighting conditions, exhibit a low-rank structure.
- providing a framework for intensity inpainting of geometrically filled-in missing regions (obtained using [20]) using the learned compact subspace.
- introducing context-based inpainting as against [6], [8] which fill-in the missing region using the best matching neighbouring patch.

2. Related Work

Inpainting of missing data is an active research topic, explored mainly along two strategies - the first one considers the image space itself to constrain the span of the data in the missing region, while the second line of research assumes that the missing region belongs to the space of images of the same scene or semantically similar scene.

PDE-based techniques such as [3] perform structure inpainting by assuming local smoothness and interpolating the available data along the hole boundary into the hole. It was used successfully to fill-in small geometric holes with no texture. Non-parametric sampling based methods like

[9] have shown success in inpainting repeating texture patterns. Building upon these methods, [6], [7], [8] combined geometry and texture inpainting by treating them separately and finally fusing them together. For the inpainting step, exemplar patches from the neighbourhood of the same image were used to constrain the appearance space of the missing region. The work in [25] improved upon [8] and modeled a patch containing damaged pixels as belonging to a span of a sparse set of atoms from an image dictionary, built using exemplars from the known regions of the image. In [12], Tensor Voting (TV)-based image segmentation is proposed to yield boundaries dividing salient regions followed by TV-based colour inference in each salient region overlapping the missing region. Inpainting local repetitive structures by promoting blockwise low-rankness is proposed in [16]. Multiscale graph cuts-based inpainting [15] formulate the problem as that of discrete global energy minimization to enforce structure and texture consistency. In [13], a search for best matching patch is employed in the feature space. However, all these methods restrict the missing region data to lie in the image space itself, and as noted in [10], an image space can not span the large dimensionality of all possible missing textures. This can lead to artifacts in the resulting image.

The second line of research mentioned earlier benefits from easy availability of plenty of images of famous tourist spots and monuments on several image sharing websites. These methods have been shown to work quite well even though the exemplar images may not be at the same scale or of the same illumination as the damaged image. In [10], filling-in of the missing region is performed using best matching patches from across the set of semantically similar scenes. Though inherently context-based because of the use of scene matching descriptors, the seamlessly blended results produced may not be true to the actual scene as no constraints are put on the type of scene exemplars used. The work in [22] used images of the same scene to accurately fill-in the missing region by finding a set of homographies that geometrically and photometrically register the damaged image and its exemplars.

3. Methodology

Following the second line of research, in this paper, we propose intensity filling-in for large damaged regions found in 3D models of archaeological structures using exemplar patches derived from internet images of other archaeological structures. The missing region is modeled as gross pixel corruption and filling-in of this region is performed using a learned low-dimensional subspace.

We would like to add that geometric reconstruction of large holes in damaged 3D models has already been shown in [20]. The resulting geometrically inpainted 3D models are used as an input for our work.

3.1. A low-rank matrix model

Claim: *Image patches corresponding to arbitrarily oriented planar regions, found locally in several Lambertian surfaces, with possibly different albedo and uncontrolled lighting conditions, lie in a low-dimensional subspace.*

Proof: Consider a set of images $\{I_i\}$, $i = \{1, 2, \dots, k\}$, corresponding to k 3D surfaces $\{X_i\}$, where the number of pixels in each I_i equals m . Let $\text{vec}(I_i)$ represent an operation that lexicographically orders I_i . Rearranging $\{\text{vec}(I_i)\}$ as column vectors of a matrix A , we have

$$A = \begin{bmatrix} \text{vec}(I_1) & \text{vec}(I_2) & \cdots & \text{vec}(I_k) \end{bmatrix} \quad (1)$$

$$= \begin{bmatrix} I_1(1) & I_2(1) & \cdots & I_k(1) \\ I_1(2) & I_2(2) & \cdots & I_k(2) \\ \vdots & \vdots & \ddots & \vdots \\ I_1(m) & I_2(m) & \cdots & I_k(m) \end{bmatrix} \quad (2)$$

where $A \in \mathbb{R}^{m \times k}$ and $I_i(j)$ is the j^{th} element of $\text{vec}(I_i)$, $j = \{1, 2, \dots, m\}$. Assuming Lambertian image formation model, the intensity at $I_i(j)$ corresponding to a surface point $\mathbf{p} = [x \ y \ z]^T$ on X_i is given by

$$I_i(j) = \rho_{ij} \mathbf{n}_{ij}^T \mathbf{s}_i \quad (3)$$

where ρ_{ij} and $\mathbf{n}_{ij} \in \mathbb{R}^3$ are the albedo and normal direction at \mathbf{p} , respectively, and $\mathbf{s}_i \in \mathbb{R}^3$ is the direction of a distant source of light. The assumption of a single distant source of light is for the sake of simplicity. This constraint will be relaxed subsequently. Substituting (3) in (2), we get

$$A = \begin{bmatrix} \rho_{11} \mathbf{n}_{11}^T \mathbf{s}_1 & \rho_{21} \mathbf{n}_{21}^T \mathbf{s}_2 & \cdots & \rho_{k1} \mathbf{n}_{k1}^T \mathbf{s}_k \\ \rho_{12} \mathbf{n}_{12}^T \mathbf{s}_1 & \rho_{22} \mathbf{n}_{22}^T \mathbf{s}_2 & \cdots & \rho_{k2} \mathbf{n}_{k2}^T \mathbf{s}_k \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1m} \mathbf{n}_{1m}^T \mathbf{s}_1 & \rho_{2m} \mathbf{n}_{2m}^T \mathbf{s}_2 & \cdots & \rho_{km} \mathbf{n}_{km}^T \mathbf{s}_k \end{bmatrix} \quad (4)$$

Equivalently, we can write

$$\rho_{ij} \mathbf{n}_{ij}^T \mathbf{s}_i = \rho_{ij} (R_i \mathbf{n}_j)^T \mathbf{s}_i = \rho_{ij} \mathbf{n}_j^T R_i^T \mathbf{s}_i, \quad \forall i, j$$

i.e., if a surface with lexicographically ordered set of normals given by $[\mathbf{n}_1 \ \mathbf{n}_2 \ \cdots \ \mathbf{n}_m]^T$ is chosen as a reference, R_i is the rotation matrix that maps $[\mathbf{n}_1 \ \mathbf{n}_2 \ \cdots \ \mathbf{n}_m]^T$ to the lexicographically ordered set of normals corresponding to an arbitrary rotation of such a surface. Let $\{I_i\}$ be the images of such surfaces taken from several images. Thus, we obtain

$$A = \begin{bmatrix} \rho_{11} \mathbf{n}_1^T R_1^T \mathbf{s}_1 & \rho_{21} \mathbf{n}_1^T R_2^T \mathbf{s}_2 & \cdots & \rho_{k1} \mathbf{n}_1^T R_k^T \mathbf{s}_k \\ \rho_{12} \mathbf{n}_2^T R_1^T \mathbf{s}_1 & \rho_{22} \mathbf{n}_2^T R_2^T \mathbf{s}_2 & \cdots & \rho_{k2} \mathbf{n}_2^T R_k^T \mathbf{s}_k \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1m} \mathbf{n}_m^T R_1^T \mathbf{s}_1 & \rho_{2m} \mathbf{n}_m^T R_2^T \mathbf{s}_2 & \cdots & \rho_{km} \mathbf{n}_m^T R_k^T \mathbf{s}_k \end{bmatrix} \quad (5)$$

If the images of the surfaces considered are of a sufficiently local region, the albedo in such a region can be assumed to

be constant, i.e., $\rho_{ij} = \rho_i, \forall i, j$. Such an assumption holds for the appearance space of monolithic stone carvings. Let us collect the normals and albedo terms into large sparse matrices N and D respectively, such that

$$N = \begin{bmatrix} \mathbf{n}_1^T & \cdots & & \\ & \mathbf{n}_2^T & \cdots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \cdots & \mathbf{n}_m^T \end{bmatrix}, D = \begin{bmatrix} \rho_1 & \cdots & & \\ & \rho_2 & \cdots & \\ \vdots & \vdots & \ddots & \vdots \\ & & \cdots & \rho_k \end{bmatrix}$$

Consequently, we can factorize A in (5) as

$$A = N \begin{bmatrix} R_1^T \mathbf{s}_1 & R_2^T \mathbf{s}_2 & \cdots & R_k^T \mathbf{s}_k \\ R_1^T \mathbf{s}_1 & R_2^T \mathbf{s}_2 & \cdots & R_k^T \mathbf{s}_k \\ \cdots & \cdots & \ddots & \cdots \\ R_1^T \mathbf{s}_1 & R_2^T \mathbf{s}_2 & \cdots & R_k^T \mathbf{s}_k \end{bmatrix} D \quad (6)$$

$$= N \quad \quad \quad S \quad \quad \quad D$$

where $D \in \mathbb{R}^{k \times k}$, $N \in \mathbb{R}^{m \times 3m}$, and $S \in \mathbb{R}^{3m \times k}$. Since S is of a form such that the first three rows are repeated $(m - 1)$ times, S has atmost 3 linearly independent rows, while D and N are full rank matrices. Analyzing the rank of the matrix A , given that m is the number of pixels considered in each I_i , k is the number of exemplars used and $3 < m < k$, we get

$$\begin{aligned} \text{rank}(A) &= \min\{\text{rank}(N), \text{rank}(S), \text{rank}(D)\} \\ &= \min\{m, \text{rank}(S), k\} \end{aligned}$$

$$\text{Therefore, } \text{rank}(A) \leq 3 \quad (7)$$

If we consider arbitrarily rotated locally planar regions as the underlying surfaces in the above analysis, only the matrix N changes ($\mathbf{n}_1 = \mathbf{n}_2 = \cdots = \mathbf{n}_m$), while the bound on $\text{rank}(A)$ remains the same. Image patches from planar surface regions with no geometry (for example, walls of archaeological structures) serve as ideal candidates to provide such an exemplar, and can easily be found in several internet images of archaeological structures.

3.1.1 Complex lighting scenario

We now extend the low-rank result derived in the section 3.1 to the situation of multiple light sources. It is well-known that the appearance space of a *single* Lambertian surface lies in a 9D space [5], [19]. Using the forward rendering model of [18], the irradiance $E(\mathbf{n})$ at a surface point \mathbf{p} with surface normal $\mathbf{n} = [n_x \ n_y \ n_z]^T$ is given by

$$\begin{aligned} E(\mathbf{n}) &= c_1 L_{22} (n_x^2 - n_y^2) + c_3 L_{20} n_z^2 + c_4 L_{00} - c_5 L_{20} \\ &\quad + 2c_1 (L_{2-2} n_x n_y + L_{21} n_x n_z + L_{2-1} n_y n_z) \\ &\quad + 2c_2 (L_{11} n_x + L_{1-1} n_y + L_{10} n_z) \end{aligned} \quad (8)$$

where L_{lm} ($l = \{0, 1, 2\}$ and $-l \leq m \leq l$) are the first 9 spherical harmonic co-efficients and c_j , $j = \{1, 2, \dots, 5\}$ are

constants as defined in [18]. The image intensity at a pixel j corresponding to a surface point \mathbf{p} is given by

$$I_i(j) = \rho_{ij} E(\mathbf{n}_{ij}) \quad (9)$$

Rearranging (8) into dot-product form by collecting the spherical harmonic co-efficients into \mathbf{L}_i and the remaining terms into the vector \mathbf{c}_{ij} , and substituting in (9) yields

$$I_i(j) = \rho_{ij} \mathbf{c}_{ij}^T \mathbf{L}_i \quad (10)$$

Since (3) and (10) are of the same form, the proof in section 3.1 can be repeated by making appropriate changes to the matrices N , S and D , the rank of which now changes to m , 9 and k respectively. Again, from (7) it follows that

$$\begin{aligned} \text{rank}(A) &= \min\{m, \text{rank}(S), k\} \\ \text{Therefore, rank}(A) &\leq 9 \end{aligned} \quad (11)$$

Thus, regardless of the choice of simple (single light source) or complex (multiple light sources) lighting model, we have shown that a set of local planar patches taken from different Lambertian surfaces indeed lies in a low-dimensional subspace.

3.1.2 Differences with [5], [19] and [24]

Our proof for (7) begins by assuming the well-known Lambertian image formation model, for which it has been shown in [5], [19] and [24] that the image space for a *single* 3D surface observed under varied lighting conditions lies in a compact subspace. Building upon that inference, this paper proposes an interesting and hitherto unknown conclusion that even local patches corresponding to arbitrarily oriented planar underlying geometry, with different albedo and illumination conditions, taken from images of *several* 3D surfaces also exhibit a low-rank matrix structure. A similar inference is neither implicitly obvious using the formulation of [5], [19] or [24] nor has it been explicitly established in the literature. Furthermore, deviating from the factorization used in [24], neither of the matrices N , S or D need to be known. Also the low-rank matrix result (7), (11) is used to inpaint missing regions in local patches as against surface normal estimation performed in [24].

3.2. Inpainting holes using the low-rank model

Let us assume the availability of $(k - 1)$ images $\{I_i\}$ of arbitrarily oriented planar surfaces with different albedos and illumination conditions as explained in section 3.1. The goal now is to estimate the missing regions in a locally planar patch P (Fig. 1(c)), which is of a size equal to that of I_i , and which partially overlaps with both the broken and undamaged regions. From (7) and (11), such a patch would form a low-rank matrix model along with the other $(k - 1)$ images.

Modeling the pixels in the missing region of the patch as gross pixel corruptions, and these corruptions as additive error, the corrupted matrix A_{obs} can be decomposed as

$$A_{obs} = A + E \quad (12)$$

where A_{obs} is the matrix of the above k images arranged as column vectors, A is the required low-rank matrix, and E is the matrix containing additive errors. Since only the k^{th} column in A_{obs} is corrupted, E turns out to be sparse. Thus, the problem reduces to estimating the following

$$\begin{aligned} \{\hat{A}, \hat{E}\} &= \arg \min_{A, E} (\text{rank}(A) + \gamma \|E\|_0) \\ &\text{s.t. } A_{obs} = A + E \end{aligned} \quad (13)$$

where $\text{rank}(\cdot)$ is a function that calculates the rank of the given matrix and $\|\cdot\|_0$ finds the l_0 -norm. Recent works in rank minimization based on convex optimization techniques [23], [24], [26], [27] have made it possible to robustly recover low-rank matrices even in the presence of significant amounts of corruption. These works state that when both the rank of the matrix to be recovered (A) and the number of errors in E are not too high, the $\text{rank}(\cdot)$ function and $\|\cdot\|_0$ function in (13) can be replaced by the nuclear norm (i.e., the sum of the singular values) and l_1 -norm (i.e., the $\|\cdot\|_1$ function), respectively. To solve (13), we use the Augmented Lagrange Multiplier (ALM) method, the source code of which has been made available by the authors of [27]. This optimization procedure enables search for the robust bases, a linear combination of which best explains the missing entries in the k^{th} column of A_{obs} to yield the low rank matrix A .

Algorithm 1 Intensity inpainting of large missing regions

Input: (a) 3D model with geometrically inpainted damaged region, (b) database of $(k - 1)$ exemplars $\{X_i\}$

Output: Intensity inpainted hole region of the 3D model

- 1: $A_{obs} = [\text{vec}(X_1) \mid \text{vec}(X_2) \mid \dots \mid \text{vec}(X_{k-1})]$
 - 2: masked hole region
 - 3: **while** hole exists **do**
 - 4: $\{h_j\} \leftarrow$ current boundary of the hole region
 - 5: **for** $j = 1$ to No. of elements in $\{h_j\}$ **do**
 - 6: $\{P\} \leftarrow$ overlapping patches containing h_j
 - 7: Append $\text{vec}(P)$ as k^{th} column of A_{obs}
 - 8: $\{\hat{A}, \hat{E}\} = \arg \min_{A, E} (\text{rank}(A) + \gamma \|E\|_0)$
s.t. $A_{obs} = A + E$
 - 9: k^{th} column of \hat{A} has the inpainted value for h_j
 - 10: Repeat steps 6 to 9 for each patch containing h_j
 - 11: Unmask h_j
 - 12: **end for**
 - 13: Average the estimates from each overlapping patch
 - 14: **end while**
-

Algorithm 1 details the steps followed for the implementation of the inpainting procedure using the methodology described above. In a manner similar to [2], overlapping patches are considered and contribution from each patch is averaged to yield the final inpainted value at each pixel in the hole region.

4. Experimental Results

In this section, we demonstrate the effectiveness of the proposed algorithm and also provide comparisons with related works both quantitatively (synthetic data) and qualitatively (real examples). Comparisons are shown against the works [8], [10] and [13] because the source code of their implementation was freely available. Although [22] is related, we do not compare with it since it makes use of the same scene exemplar for inpainting. The proposed method addresses a scenario where a ‘completed’ same scene may not be available at all, and thus a comparison with [22] would be unfair. Comparisons with CSH-based inpainting [13] is provided as it has been shown to perform better than Patch-Match [4].

As detailed in section 3.1, a corpus of reliable exemplars needs to be created to proceed with the proposed low-rank model-based inpainting. Also, the same examples database should be applicable for all the inpainting scenarios being considered. To ensure good span across the appearance space, a variety of images of archaeological structures available on online image sharing websites are made use of. Exemplar image patches from surface regions with no geometry corresponding to roughly planar regions are manually cropped and stored, and the same database is used across all the experiments. The patch size chosen in this work is 20×20 as it ensures patch planarity for inpainting and provides the best trade-off between ensuring overcompleteness of A (given that the number of patches used is 1400), and speed of execution. It is to be noted that small patches from arbitrarily oriented planar regions can be expected to have some rotation times the similar set of normals (5). However, if this gets violated, $\text{rank}(A)$ will go up, reflecting in an increase in reconstruction error for the known region in patch P , leading to artifacts. From the results, it will be evident that this is an unlikely event. Also, empirically, the $\text{rank}(A)$ turned out to be close to 9 for almost all of the inpainted patches, leading to the conclusion that complex lighting source model well-represents uncontrolled illumination scenarios.

4.1. Synthetic data

Implementing the proposed algorithm on synthetic data allows for quantitative comparison with other works proposed in the literature. Evaluating against the ground truth, the reconstruction error is determined by two standard metrics used commonly in the colour correction literature,

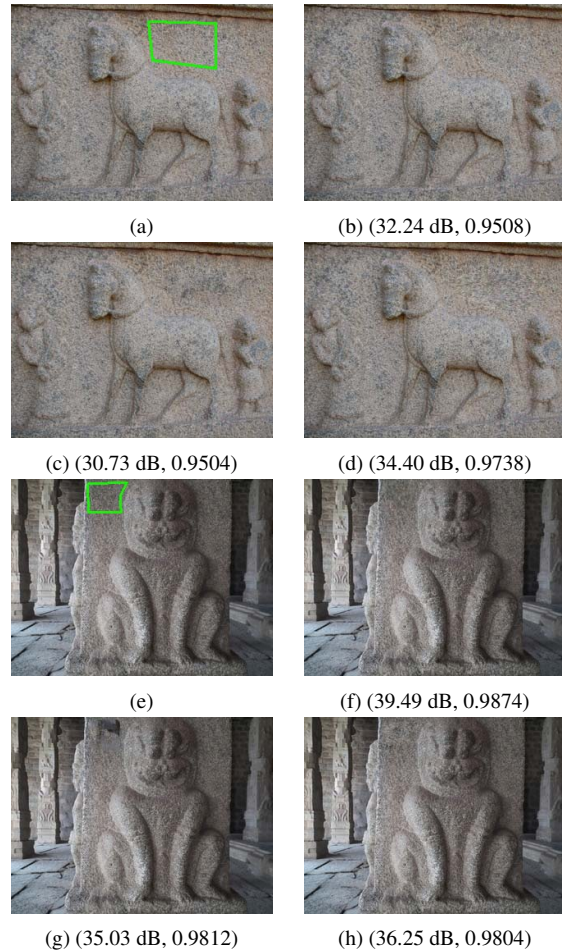


Figure 2. Intensity inpainting comparison for a synthetically generated ‘missing’ region with no geometry. (PSNR, SSIM) values are given in parentheses. (a, e) Original images, (b, f) results using the proposed method, (c, g) results using [8], (d, h) results using [10]. The masked region is shown bounded in green.

namely PSNR and SSIM [21].

In the first example (Figs. 2 (a), (e)), a known textured region with no geometry is masked (shown bounded in green) and treated as ‘missing’ region. Since the marked region in Fig. 2 (a) does not contain any significant geometry or texture variations, there is no distinct advantage provided by the proposed method (Fig. 2 (b)). On the other hand, the marked region in Fig. 2 (e) contains varying texture and is near to the boundary of the foreground object and the background. In this case, performance using the proposed method (Fig. 2 (f)) turns out to be better than the output provided by [10] (Fig. 2 (h)). In both the cases, the proposed method outperforms [8] (Figs. 2 (c), (g)), which can be attributed to the usage of context information in our scheme.

The second example (Fig. 3) considers a more interesting synthetic scenario, where a missing region is created

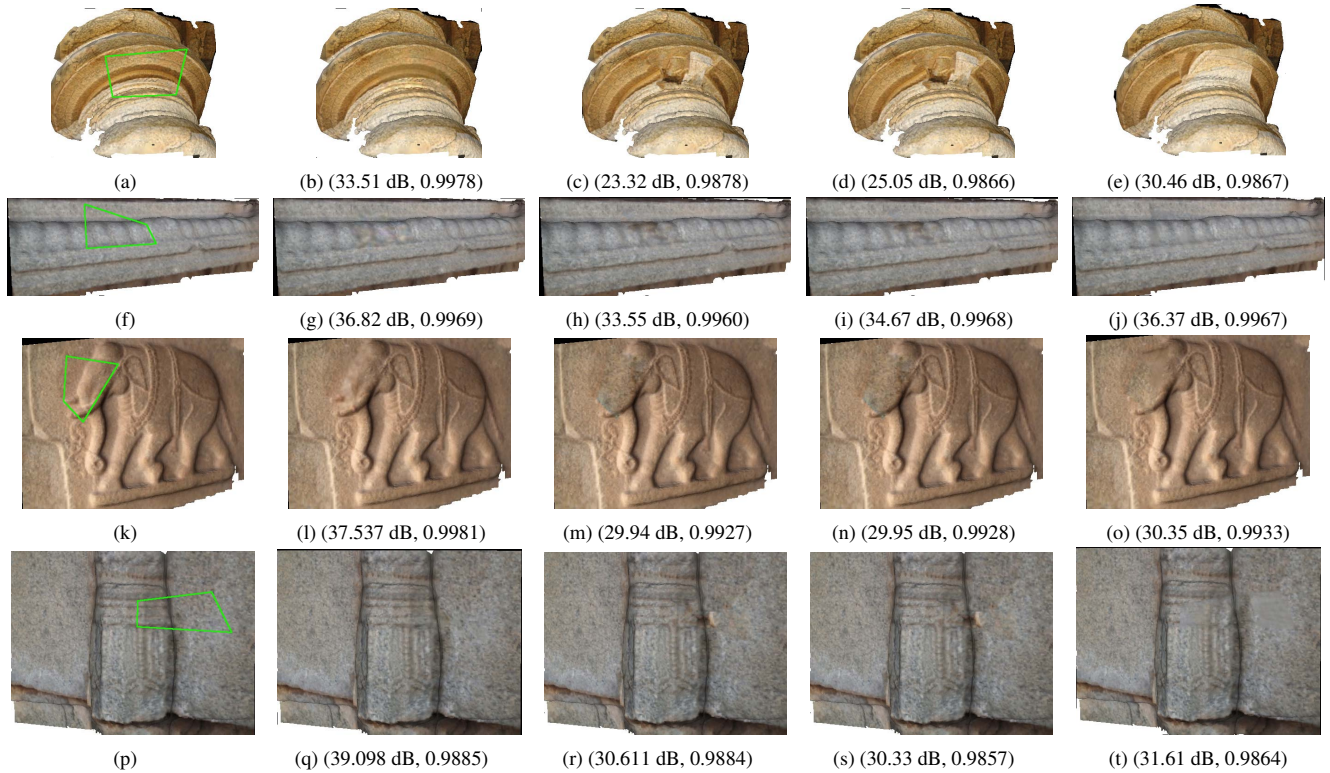


Figure 3. Inpainting of synthetically created ‘missing’ regions in 3D models. (PSNR, SSIM) pair are given in parentheses for each result. (a, f, k, p) Original 3D model, (b, g, l, q) results using the proposed method, (c, h, m, r) results using [8], (d, i, n, s) results using [10], and (e, j, o, t) results using [13].

so as to include significant 3D geometry variations (for example, presence of curved areas). A visual analysis of the results delivered by the proposed method reveals very few artifacts. Quantitative analysis in the form of PSNR and SSIM values too serve to corroborate the performance of our method.

It is to be noted that although the examples shown in Figs. 2 and 3 consider a synthetically generated ‘missing’ region, the scenarios considered (especially Fig. 3) are quite complex with a considerable chunk of region marked as ‘missing’. Such a hole-region in no way simplifies the inpainting problem.

Discussions on comparative results : Neighbourhood or patch-based methods [8], [13] have been shown to work well if the known region in the image contains significant data variations to correctly fill-in the missing region. However, as noted in [10], a single image space does not span the large dimensionality of all possible missing textures. Also, the self-similar examples used in [20] typically come from multiple archaeological sites and exhibit different material properties, rendering search for best matching patch useless. All these factors lead to an incorrect filling-in using neighbourhood or patch-based methods. In comparison, our

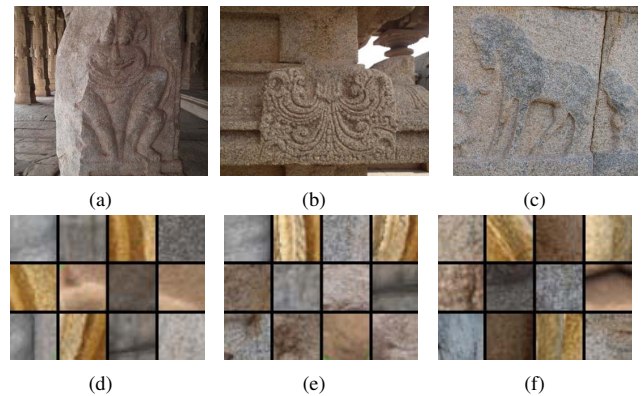


Figure 4. (a, b, c) One of the views of the three real examples considered with large naturally existing damaged regions. (d, e, f) Sample patches from across many examples showing texture variations.

method learns robust basis of a low-dimensional subspace from varied exemplar patches obtained from a large set of images to correctly model the missing region in a given patch.



Figure 5. Geometric and photometric inpainting of real examples. (a, f, k) Original 3D model, (b, g, l) results using the proposed method, (c, h, m) results using [8], (d, i, n) results using [10], and (e, j, o) results using [13].

4.2. Real objects

The scenario of naturally existing large missing regions in real world structures is considered next. Such structures can be easily found, especially at archaeological sites. All the examples considered in this work are of sculptures of archaeological interest (Figs. 4 (a), (b), (c)). Due to absence of ground truth information, a quantitative error analysis is not possible for this scenario and the results can only be judged qualitatively.

The 1st row of Fig. 5 shows geometric and photometric inpainting result for the broken regions of the 3D model of a mythical lion stone carving. The results using the proposed method are shown along with those of [8] and [10]. The second real example considered is that of a ‘lion head’ (2nd row of Fig. 5) where the entire head region was missing. Another example considered is that of a horse carving (3rd row of Fig. 5) where the body of the horse was

severely damaged. The absence of any artifacts or any visible seam at the junction of reconstructed and undamaged regions (Figs. 5 (b), (g), (l)) indicates the effectiveness of the proposed method.

Discussions on comparative results : Since the region to be inpainted is known to belong to the appearance space of sample patches forming the low-rank structure, the results using the proposed method are visually more salient in comparison to other works which pick up best matches without restricting the context, leading to visible artifacts (3rd, 4th and 5th column of Fig. 5). Though the context-based inpainting followed in the proposed method may appear to be restrictive, the variety of textures inpainted (Figs. 4 (d), (e), (f)) illustrates the effectiveness of our method. The availability of large exemplars from internet images greatly simplifies the search for exemplar patches ‘describing’ the context, and simplifies the implementation process

as well.

5. Conclusions

The proposed work allows for visualizing, with saliently inpainted intensity values, large broken regions in 3D models which have been geometrically hole-filled. Such a combined 3D geometric and photometric methodology would allow for creating visually pleasing digital archives of heritage sites, and would also be a natural addition to applications aiming to provide a ‘Virtual Tour’ of archaeological sites. The main contribution of this work includes proposing context-based inpainting which ensures a visually artifact-free intensity filling-in and inpainting of the missing region using a learned low-rank subspace formed by exemplar patches taken from internet images. The high quality and visually salient intensity inpainted hole regions in the resulting 3D models adequately demonstrate the effectiveness of the proposed method.

References

- [1] Draft charter on the preservation of the digital heritage. In *UNESCO General Conference 32nd session*, 2003. 1
- [2] M. Aharon, M. Elad, and A. Bruckstein. K-svd: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Sig. Proc.*, 54(11):4311–4322, Nov. 2006. 5
- [3] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera. Filling-in by joint interpolation of vector fields and gray levels. *IEEE Trans. on Imag. Proc.*, 10(8):1200–1211, Aug. 2001. 2
- [4] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. *ACM Trans. Graph.*, 28(3):24:1–24:11, July 2009. 5
- [5] R. Basri and D. W. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(2):218–233, 2003. 3, 4
- [6] M. Bertalmio, L. A. Vese, G. Sapiro, and S. Osher. Simultaneous structure and texture image inpainting. *IEEE Trans. on Imag. Proc.*, 12(8):882–889, 2003. 2
- [7] A. Bugeau, M. Bertalmio, V. Caselles, and G. Sapiro. A Comprehensive Framework for Image Inpainting. *IEEE Trans. on Imag. Proc.*, 2010. 2
- [8] A. Criminisi, P. Perez, and K. Toyama. Region filling and object removal by exemplar-based image inpainting. *IEEE Trans. on Imag. Proc.*, 13(9):1200–1212, Sept. 2004. 2, 5, 6, 7
- [9] I. Drori, D. Cohen-Or, and H. Yeshurun. Fragment-based image completion. *ACM Trans. Graph.*, 22(3), 2003. 2
- [10] J. Hays and A. A. Efros. Scene completion using millions of photographs. *ACM Trans. Graph.*, 26(3), 2007. 1, 2, 5, 6, 7
- [11] A. Hertzmann and S. M. Seitz. Example-based photometric stereo: Shape reconstruction with general, varying brdfs. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(8):1254–1264, Aug. 2005. 1
- [12] J. Jia and C.-K. Tang. Tensor voting for image correction by global and local intensity alignment. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(1):36–50, 2005. 2
- [13] S. Korman and S. Avidan. Coherency sensitive hashing. In *ICCV*, pages 1607–1614, 2011. 2, 5, 6, 7
- [14] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg, J. Shade, and D. Fulk. The digital michelangelo project: 3d scanning of large statues. In *ACM SIGGRAPH 2000*, pages 131–144, 2000. 1
- [15] Y. Liu and V. Caselles. Exemplar-based image inpainting using multiscale graph cuts. *IEEE Trans. on Imag. Proc.*, 2012. 2
- [16] S. Ono, T. Miyata, I. Yamada, and K. Yamaoka. Missing region recovery by promoting blockwise low-rankness. In *ICASSP*, pages 1281–1284, 2012. 2
- [17] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Y. Ma. Rasl: Robust alignment by sparse and low-rank decomposition for linearly correlated images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34:2233–2246, 2012. 2
- [18] R. Ramamoorthi and P. Hanrahan. An efficient representation for irradiance environment maps. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, SIGGRAPH ’01, pages 497–500, 2001. 3, 4
- [19] R. Ramamoorthi and P. Hanrahan. On the relationship between radiance and irradiance: determining the illumination from images of a convex lambertian object. *J. Opt. Soc. Am. A*, 18(10):2448–2459, Oct 2001. 3, 4
- [20] P. Sahay and A. N. Rajagopalan. Harnessing self-similarity for reconstruction of large missing regions in 3d models. In *ICPR*, pages 101–104, 2012. 1, 2, 6
- [21] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. on Imag. Proc.*, 13(4):600–612, 2004. 5
- [22] O. Whyte, J. Sivic, and A. Zisserman. Get out of my picture! internet-based inpainting. In *BMVC*, 2009. 2, 5
- [23] J. Wright, Y. Ma, and Y. Peng. Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization. In *NIPS*, 2009. 2, 4
- [24] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *ACCV*, pages 703–717, 2011. 4
- [25] Z. Xu and J. Sun. Image inpainting by patch propagation using patch sparsity. *IEEE Trans. on Imag. Proc.*, 19(5):1153–1165, May 2010. 2
- [26] Z. Zhang, A. Ganesh, X. Liang, and Y. Ma. Tilt: Transform invariant low-rank textures. *Int. J. Comput. Vision*, 99(1):1–24, 2012. 4
- [27] L. Zhouchen, M. Chen, L. Wu, and Y. Ma. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. *UIUC Tech. Report UILU-ENG-09-2215*, 2009. 2, 4