

Distortion Optimization based Image Completion from a Large Displacement View

Chunxiao Liu, Yingzhen Yang, Qunsheng Peng[†], Jin Wang, Wei Chen[‡]

State Key Lab of CAD&CG, Zhejiang University, China

Abstract

We present a new image completion method based on an additional large displacement view (LDV) of the same scene for faithfully repairing large missing regions on the target image in an automatic way. A coarse-to-fine distortion correction algorithm is proposed to minimize the perspective distortion in the corresponding parts for the common scene regions on the LDV image. First, under the assumption of a planar scene, the LDV image is warped according to a homography to generate the initial correction result. Second, the residual distortions in the common known scene regions are revealed by means of a mismatch detection mechanism and relaxed by energy optimization of overlap correspondences, with the expectations of color constancy and displacement field smoothness. The fundamental matrix for the two views is then computed based on the reliable correspondence set. Third, under the constraints of epipolar geometry, displacement field smoothness and color consistency of the neighboring pixels, the missing pixels are orderly restored according to a specially defined repairing priority function. We finally eliminate the ghost effect between the repaired region and its surroundings by Poisson image blending. Experimental results demonstrate that our method outperforms recent state-of-the-art image completion methods for repairing large missing area with complex structure information.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation - Bitmap and Framebuffer Operations; I.4.8 [Image Processing and Computer Vision]: Scene Analysis - Stereo

1. Introduction

Image completion concerns the problem of filling in the occluded or damaged regions of an image with the available information from the same image or another to generate visually plausible result. Due to the wide applications in photo editing and special effects production [CK04], image completion has received great attention in the past decade, and many methods have been brought forward [SC05].

Traditional image completion methods can be roughly classified into either partial differential equations (*PDEs*) or texture synthesis based approaches. Most of the former methods work by propagating the color of the surrounding known pixels into the missing regions. They usually work well only for small scratches, but may fail for those

with complex texture. The latter methods can produce compelling completion results for relatively large missing region with distinct texture by copying small source fragments from known regions. Nevertheless, they can hardly restore the structure information due to the following two fundamentally unreasonable assumptions. First, they assume that the missing pixels can be found in the surrounding known regions on the target image. Generally speaking, such an assumption cannot be satisfied in most cases. Therefore, the definition of search space makes it in nature an ill-posed problem [She03]. Second, once the best source fragment is found, it has to be transformed to fill in the target location. However, it is usually assumed that the scene in the fragment is planar and aligned to the image plane. All the 3D information embedded in the scene is ignored.

As can be seen, most of the existing methods are mainly single image based, and cannot work well for the large missing regions, especially for those with strong structures. To ensure a faithful restoration of the large missing regions, we

[†] Corresponding author

[‡] {liuchunxiao,yangyingzhen,peng.jwang,chenwei}@cad.zju.edu.cn

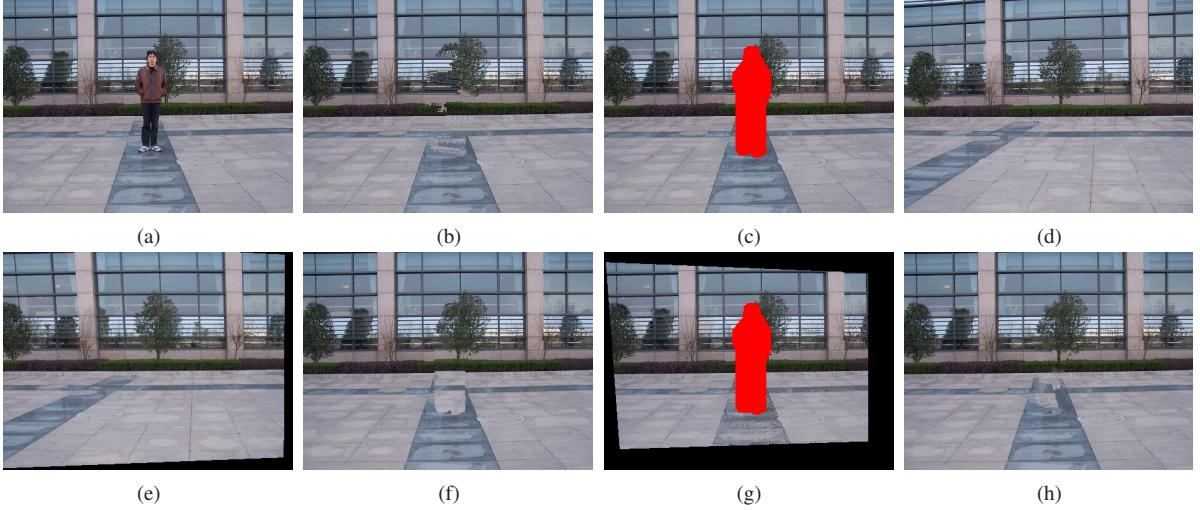


Figure 1: Human removal. (a) The target image; (b) The repairing result by texture synthesis based method [CPT04]; (c) Human removal with the occluded pixels in red; (d) The LDV image; (e) The warped LDV image; (f) Image stitching result [Sze04] with obvious mismatch on the ground; (g) Optimization result of overlap correspondences; (h) Our repairing result.

propose here a new image completion method based on an additional large displacement view (LDV) image of the same scene, which gets rid of the above two assumptions. The additional LDV view can be obtained by moving a consumer camera to keep away from the obstacles and make the previously occluded scene visible.

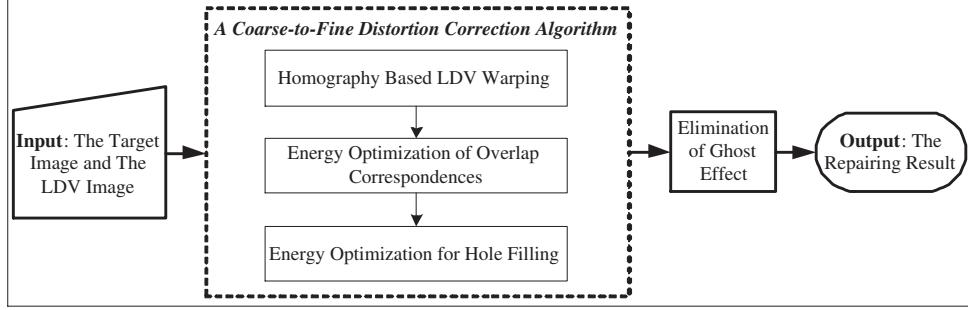
Although few methods address image completion based on views of large displacement, it provides a feasible solution for image completion of our daily photos [LGP*07], especially when there are uninterrupted crowd in some famous resorts or immovable obstacles in front of the scene of interest. The problem is therefore how to restore the occluded scene after removal of the intruders or obstacles, and generate the visually pleasing result. It would be difficult to complete the current view image by itself, especially for missing regions with complex structures. In this case, another view of the same scene with large displacement may reveal the previously occluded regions, enabling a faithful restoration of the occluded regions on the target image.

However, to achieve a faithful completion, two key challenges have to be investigated. One challenge is how to correct the perspective distortion in the corresponding parts for the common known scene regions on the LDV image. Directly using the globally warped LDV image to repair the damaged regions will result in poor result. This is because the LDV image may contain components of different depths. We then need to find better pixel correspondences over the common scene regions of the two views. A new energy optimization scheme for dense pixel correspondences of the LDV views is proposed under the constraints of color constancy and displacement field smoothness. During the course

of optimization, we set a mismatch detection mechanism to pick up the pixels which lose correspondences and adopt a dynamic increasing weighting parameter to correct them. For numerical stability, the optimization is conducted pixel-wise.

Another challenge is how to estimate the missing regions on the target image based on the rectified surrounding known scene regions. We also treat this problem as energy optimization of pixel correspondences. According to epipolar geometry, the candidate pixel on the LDV image lies on the epipolar line determined by the current pixel on the target image and a fundamental matrix [HZ00]. Accounting for the displacement field smoothness and color consistency of the neighboring pixels as two additional constraints, we derive a new energy function to predict the missing pixels. To restore the potential image structure in the missing regions, we perform an ordered restoration according to a predefined repairing priority function. To eliminate the ghost effect due to luminance difference between the target image and the LDV image, we further adopt Poisson image blending [PGB03] to generate a seamless completion.

The proposed method makes contributions in the following three aspects. First, we present a new distortion minimization approach for image completion based on a LDV image. A coarse-to-fine distortion correction algorithm is proposed to restore the missing pixels on the target image. Second, we present a new energy optimization strategy to solve the dense pixel correspondences on the LDV images, which is essentially correcting the residual distortions in the common known scene regions after applying the global warp. Third, under the constraints of epipolar geometry, dis-

**Figure 2:** The algorithm overview.

placement field smoothness and color consistency of the neighboring pixels, view consistent hole filling is achieved by a new energy optimization scheme.

The rest of the paper is organized as follows. Section 2 reviews the related work. Section 3 describes our approach of image completion in detail based on distortion optimization of the LDV image. Experimental results and analysis are given in Section 4. Section 5 concludes the whole paper and highlights future work.

2. Related work

This paper is inspired by previous work on image and video completion. We review them here.

2.1. Image completion

Since Bertalmio *et al.* [BSCB00] presented their work on image inpainting for the first time in SIGGRAPH' 2000, many methods have been brought forward. Nevertheless, most of them are single image based and can be classified as partial differential equations (*PDEs*) based methods, texture synthesis based methods, and statistics based ones.

PDEs based methods regard image completion as PDEs solving [BSCB00] or variational problems [BCV*01, CS05] by specifying the known pixels around the damaged regions as boundary conditions. The image repairing process is therefore the diffusion of the known pixels into the missing regions. Such methods work well for small scratches, but may fail for highly textured regions.

Texture synthesis based methods select the known regions on the target image as texture swatches, then perform texture synthesis to generate new image fragments for the missing regions [DCOY03, JT03, CPT04, KT06]. These methods produce satisfactory results for the textured regions, but can hardly recover the precise structure information in the large missing regions. Under the guidance of specified structure curves in the blank regions [SLJS05] or specified projective transformation in the scene [PSK06], better repairing results

are generated by interactive image completion techniques. However, they demand tiresome user interaction for the natural scene with complex structures. The most recent work in [HE07] patches up holes in images by retrieving similar image regions from a huge database of photographs, therefore the size of the image database have a great influence on the final repairing result.

Statistics based methods solve the problem of image completion by statistical analysis. Levin [LZW03] obtained the global statistical distribution based on the available part of the image by statistical learning, and found the most probable image by loopy belief propagation. EM based method [FS05] treats the problem as the estimation of the missing or damaged regions, and adopts expectation maximization (EM) algorithm for ML estimation based on sparse representation of image completion.

To the best of our knowledge, few methods focus on image completion based on views of large displacement. And, the previous work [LGP*07] for this problem is limited for the quasi-planar scenes and requires additional effort for the interactive segmentation of planar scene regions.

2.2. Video completion

Bertalmio *et al.* [BBS01] extended PDEs based image completion to video completion. It is capable of filling small textureless holes on each video frame, but unsuitable for completing large holes. Regarding video completion as a global optimization problem on texture synthesis, the methods in [WSI04] and [SMKT06] recover the missing information by direct sampling spatio-temporal patches of local structures or motion. Jia *et al.*'s method [JHM05] searches for the optimal matched fragments in the video sequences and imposes the constraints on the selected patches to maintain temporal consistency. Motion periodicity is also utilized for texture synthesis based video repairing [JTWT06]. The recent work by Patwardhan *et al.* [PSB07] segments each frame into static background and moving foreground, then conducts texture synthesis based motion completion and background inpainting in order.

As can be seen, most video completion techniques assume small camera motion between adjacent frames. They seldom consider the 3D information in the scene and just fetch the corresponding parts on adjacent frames for repairing. However, for the problem discussed in this paper, we cannot directly paste the candidate parts on the LDV image onto the target image due to the severe perspective distortion of the two views.

3. Distortion Optimization based image completion from a LDV image

In this paper, we propose an algorithm for repairing the large missing regions on the target image by employing the available information from another view with large displacement based on distortion minimization. The rectified LDV image is then used for completion purpose. Two key issues are considered, i.e., how to correct the perspective distortion in the common known scene regions on the LDV image and how to estimate the missing regions with the rectified surrounding known regions of the LDV image. A coarse-to-fine distortion correction algorithm is developed to solve the above two issues. Fig.2 illustrates the overview of our algorithm.

The following subsections will elaborate on the individual stages and provide the details of our approach.

3.1. Homography based LDV warping

In order to accelerate the convergence, we first assume that the scene in the two views is approximatively located on a 3D plane. The LDV image can therefore be warped through a homography matrix to generate a globally optimal distortion correction result in the least square sense.

After specifying the missing regions on the target image, we carry out the following steps in order:

(1)*Feature Detection*: A robust feature point extraction method, i.e. scale invariant feature transform (*SIFT*) feature detector [Low04], is employed to find enough feature points on the two views and set up their high-dimensional feature descriptors.

(2)*Feature Matching*: The approximate nearest neighbor (ANN) searching algorithm [AMN*98] is exploited to establish the feature correspondences among the detected feature points on the two views.

(3)*Homography Solving*: Some outliers may exist among the feature correspondences due to image noises. We adopt the RANSAC algorithm and the Levenberg-Marquardt algorithm [HZ00] to reject the outliers and robustly solve the homography matrix H via $p = Hp'$, i.e.

$$\lambda \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} h_0 & h_1 & h_2 \\ h_3 & h_4 & h_5 \\ h_6 & h_7 & 1 \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix},$$

where p and p' are homogeneous coordinates of the matched

feature points, λ is a homogeneous constant, h_0, \dots, h_7 are the parameters of H .

(4)*Image Warping*: With H , we warp the LDV image S onto the view of the target image T as shown in Fig.3.

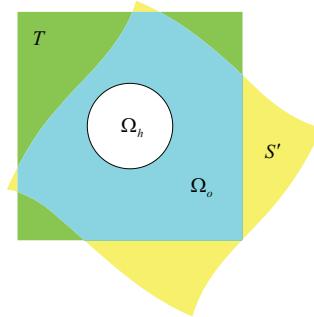


Figure 3: Homography based LDV warping. The warped LDV image S' is pasted onto the target image T , providing the initial estimate of the missing pixels in the hole Ω_h and the initial distortion correction for the common known scene region in the overlap Ω_o .

The warped LDV image S' falls on the target image T . The overlapped region $S' \cap T = \Omega_o \cup \Omega_h$ is approximately regarded as the common scene region on the two views. The part of S' which laps over the known regions on T , establishes the initial pixel correspondences for the common known scene region in the overlap Ω_o . The part of S' which covers the missing region, provides an initial estimate of the missing pixels in the hole Ω_h .

3.2. Energy optimization of overlap correspondences

Since the LDV image is warped under the rough assumption of a planar scene, there exist a large number of mismatches, i.e. residual perspective distortions, in the common scene region $\Omega_o \cup \Omega_h$ for a general scene. Directly adopting the warped LDV image S' to fill in the corresponding missing pixels would yield ugly repairing results. Because the missing region Ω_h provides no available information for correction, we can only modify the initial estimate for the missing pixels with the aid of the common known scene regions Ω_o . As a result, the residual distortions in the overlap must be further rectified in advance. We formulate it as a problem of energy optimization for pixel correspondences.

Traditional stereo matching algorithms [SS02] and optical flow algorithms [BB95] are suitable only for the dense pixel correspondences problem with short baseline. Although view morphing [SD96] deals with a similar problem as ours of establishing the pixel correspondences between two large displacement views, it requires the viewing transformation information of the two images. The most related work is the wide-baseline stereo matching method [STG03],

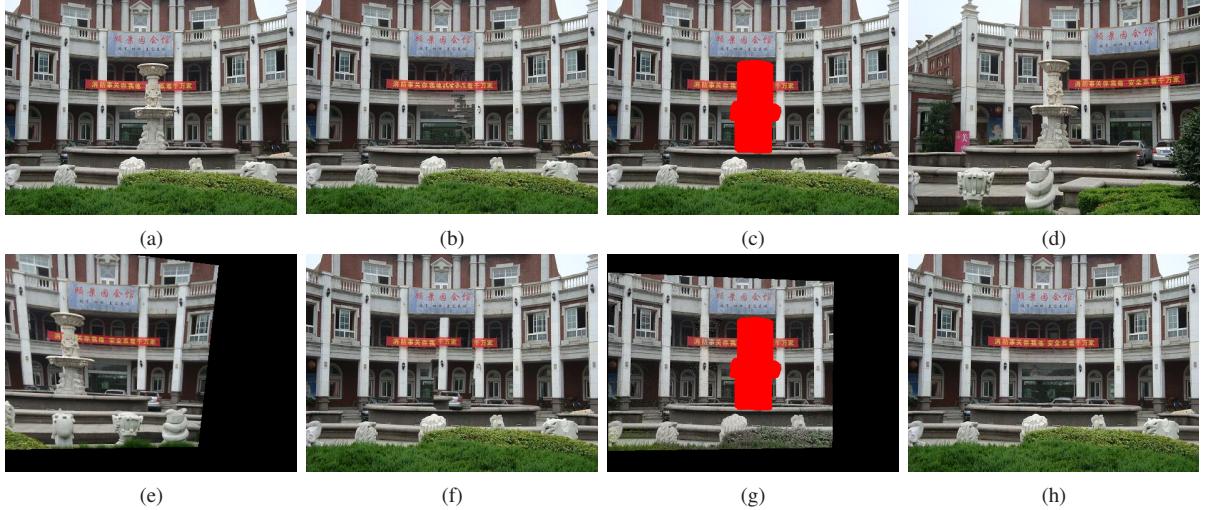


Figure 4: Object removal for the marble sculpture. (a) The target image; (b) The repairing result by texture synthesis based method [CPT04]; (c) Object removal with the occluded pixels in red; (d) The LDV image; (e) The warped LDV image; (f) Image stitching result [Sze04] with obvious mismatch on the wall and the banner; (g) Optimization result of overlap correspondences; (h) Our repairing result.

however it adopts the computationally intensive structure-from-motion as a processing stage and needs multiple views rather than two views for a stable solution. With a mismatch detection mechanism and a dynamic weighting parameter, we propose a new optimization scheme to generate reliable pixel correspondences for the LDV views.

Let (p, p') denote a corresponding pixel pair on the target image p and the LDV image p' . Let N_p represent p 's 4-connected neighbors in Ω_o , and $\langle p, q \rangle$ is a pixel pair such that $q \in N_p$. We model the common known scene region as a Markov Random Field (MRF), a pixel can then be determined by its neighbors. With the expectations of color constancy for the corresponding pixels and displacement field smoothness within neighboring pixels, we define the energy function as

$$E = \sum_{p \in \Omega_o} E_c(p) + \lambda \sum_{\langle p, q \rangle \in \Omega_o} E_s(p, q), \quad (1)$$

where $E_c(p)$ is the color constancy term, $E_s(p, q)$ is the displacement field smoothness term. λ is a dynamic weighting parameter used to balance the influences of $E_c(p)$ and $E_s(p, q)$.

The color constancy term ensures that the corresponding pixel color $C'_{p'}$ on the LDV image agrees with C_p on the target image. It is expressed as the squared difference of the corresponding pixel colors, i.e.,

$$E_c(p) = (C'_{p'} - C_p)^2. \quad (2)$$

In addition, the displacement field is expected to be smooth. Thereby, the smoothness term penalizes the inconsistent displacement changes between two neighboring pixels p and q . It is usually formulated as the squared difference

[BB95] between the corresponding pixel coordinates p' and q' . However, such an expression strongly suppresses the discontinuous motion edge and results in over-smoothness. In order to relax the penalty for the large displacement change, we instead adopt a less increasing function, i.e. Huber function $\rho(x)$. The displacement field smoothness term adopts the following form

$$E_s(p, q) = \rho(p' - q'). \quad (3)$$

thereinto

$$\rho(x) = \begin{cases} \frac{x^2}{2\delta|x| - \delta^2}, & |x| \leq \delta \\ \frac{x^2}{2\delta|x| - \delta^2}, & |x| > \delta \end{cases}.$$

Due to the irregularity of the image data and the large scale unknowns, simultaneous optimization of displacement vectors for all pixels in the overlap not only often traps in the unstable solution with local minimum but also is computationally intensive. Inspired by the work in [GS98], a relaxed optimization strategy is adopted here to solve this problem, i.e., we only optimize the displacement vector of one pixel each time and the others are fixed. Considering the energy function Eq.(1), we can see that only a few terms in it vary with the change of the displacement vector for one pixel p . Let $(X_{i,j}, Y_{i,j})$ denote the corresponding coordinates of p' on the LDV image. (i, j) is the coordinates of p on the target image. We can formulate the energy function for one pixel p as follows

$$\begin{aligned} E(p) = & (C'_{p'} - C_p)^2 + \lambda[\rho(X_{i,j} - X_{i-1,j} - 1) + \\ & \rho(X_{i,j} - X_{i,j-1}) + \rho(X_{i+1,j} - X_{i,j} - 1) + \\ & \rho(X_{i,j+1} - X_{i,j}) + \rho(Y_{i,j} - Y_{i-1,j}) + \\ & \rho(Y_{i,j} - Y_{i,j-1} - 1) + \rho(Y_{i+1,j} - Y_{i,j}) + \end{aligned}$$

$$\rho(Y_{i,j+1} - Y_{i,j} - 1)]. \quad (4)$$

During the optimization of the displacement vector for each pixel, the conjugate gradient method is applied to find the minimum of Eq.(4). The detailed optimization strategy for overlap correspondences is summarized in Table.1.

Table 1: Energy Optimization of Overlap Correspondences.

-
1. Initialization.
Initialize p' with inverse homography H^{-1} , i.e., $p' \sim H^{-1} p, \forall p \in \Omega_o$.
 2. Weight setup.
Let λ_0 be the initial value of λ .
 3. Energy function minimization.
Update p' by minimizing Eq.(4) using the conjugate gradient method, $\forall p \in \Omega_o$.
 4. Mismatch detection.
If $DisplaceTerm(p) > \theta_{dt}$ &&
 $(ColorSimilarity(p) > \theta_{cs} \parallel$
 $ColorDiffer(p) > \theta_{cd})$,
push p into the mismatch set M_m ;
else push p into the good correspondence set M_g .
 5. Optimization loops. $\forall p \in M_m$.
 - Increase the value of λ and repeat steps 3 ~ 4.
 - If p doesn't conform with the mismatch conditions,
move p from M_m to M_g .
 6. Termination conditions.
If $M_m = \emptyset$ or none is removed from M_m in the successive optimization loops, exit.
-

In Table.1, a strict mismatch detection mechanism is adopted to reject those pixels which lose correspondences due to occlusion or self-occlusion. Specifically, $DisplaceTerm(p)$ represents the displacement field smoothness term related to p in Eq.(4). $ColorSimilarity(p)$ denotes the sum of absolute color differences between the 8-neighborhood pixels of p on the target image and those of the corresponding pixel p' on the LDV image. $ColorDiffer(p)$ gives the absolute color difference between p and p' . During optimization, the dynamic weighting parameter λ is empirically initialized in the range of $0.05 \sim 0.5$ and increased by 0.5 after each optimization loop to strengthen the constraint of displacement field smoothness term for those lost pixel correspondences. Further, we have adopted an image pyramid based multi-scale procedure [OABB85] for the basic approach in Table.1 to avoid its convergence into a local minimum.

With sufficient reliable pixel correspondences $(p, p') \in M_g$ in the overlap, we can estimate a fine fundamental matrix F as a geometric constraint for the two views via the

normalized 8-point algorithm [HZ00]. Similar to the estimation of the homography matrix, we also use the RANSAC algorithm and the Levenberg-Marquardt algorithm to reject the outliers and robustly compute F satisfying $p'^T F p = 0$, i.e.,

$$\left(\begin{array}{ccc} X_{i,j} & Y_{i,j} & 1 \end{array} \right) \cdot \left(\begin{array}{ccc} f_0 & f_1 & f_2 \\ f_3 & f_4 & f_5 \\ f_6 & f_7 & 1 \end{array} \right) \cdot \left(\begin{array}{c} i \\ j \\ 1 \end{array} \right) = 0,$$

where f_0, \dots, f_7 are the eight unknown parameters of F .

3.3. Energy optimization for hole filling

Hole filling is regarded as a pixel-wise energy optimization problem: given the reliable correspondence set M_g in the common known scene regions, it aims to estimate the corresponding pixel p' on the LDV image for the missing pixel $p \in \Omega_h$ with its initial value obtained from the homography based LDV warping. To find a valid solution, basically there are three a priori expectations. First, according to epipolar geometry [HZ00], the corresponding pixel p' on the LDV image must fall on the epipolar line of p . In other words, p and p' satisfy the epipolar constraint constructed by the fundamental matrix F , i.e., $p'^T F p = 0$. Second, the displacement field around and within the hole should be smooth. Third, the color distribution in the local neighborhood of p on the target image should be consistent with that of p' on the LDV image.

Let $\partial\Omega_h$ denote the boundary of the hole, $\delta\Omega_h = \{p \in T \setminus \Omega_h : N_p \cap \Omega_h \neq \emptyset\}$ be the surrounding known pixels in the overlap. NB_p is the 3×3 image fragment centered at p . Considering all afore-mentioned constraints, we define the energy function for the missing pixel p to be repaired as follows

$$E(p) = \lambda_c E_c(p) + \lambda_e E_e(p) + \lambda_s E_s(p), \quad (5)$$

where $E_c(p)$ is the color consistency energy, $E_e(p)$ is the epipolar constraint energy, and $E_s(p)$ is the displacement field smoothness energy. λ_c , λ_e and λ_s are three weighted parameters for the balance purpose.

In Eq.(5), $E_c(p)$ is formulated as the sum of the squared color differences between $p_i \in NB_p$ on the target image and $p'_i \in NB_{p'}$ on the LDV image, i.e.,

$$E_c(p) = \sum_{p_i \in NB_p} (C'_{p'_i} - C_{p_i})^2. \quad (6)$$

$E_e(p)$ is expressed as the squared epipolar geometry errors for p and p' , i.e.,

$$E_e(p) = d^2(p', F p) + d^2(p, F^T p'). \quad (7)$$

thereinto $d(x, l)$ represents the distance from a point x to a line l .

Let R denote the set of repaired pixels in Ω_h . $E_s(p)$ is

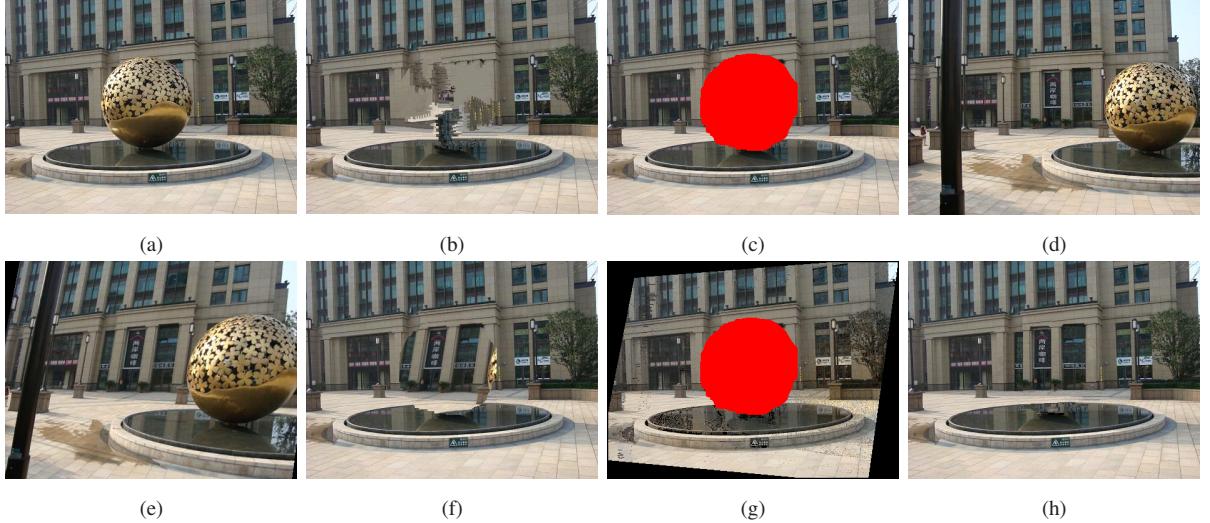


Figure 5: Object removal for the spherical statue. (a) The target image; (b) The repairing result by texture synthesis based method [CPT04]; (c) Object removal with the occluded pixels in red; (d) The LDV image; (e) The warped LDV image; (f) Image stitching result [Sze04] with obvious mismatch on the building and the pond; (g) Optimization result of overlap correspondences; (h) Our repairing result.

denoted as the sum of the squared displacement differences between p and $q \in N_p \cap ((\delta\Omega_h \cap M_g) \cup R)$, i.e.

$$E_s(p) = \sum_{(p,q) \in (\delta\Omega_h \cap M_g) \cup R} (p' - q')^2. \quad (8)$$

The repairing process for the missing pixels is conducted pixel-wise, and starts from the boundary pixels of the hole B_h with high repairing priority. For the sake of reliable hole filling and structure preserving, we define the repairing priority for $p \in B_h$ as

$$P(p) = C(p) * D(p) * S(p),$$

where

$$C(p) = \sum_{q \in NB_p} cw(q)/8$$

is the color confidence term that represents the reliable color information contained in p 's 8-neighborhood NB_p , where $cw(q)$ is the color weight of p 's 8-neighborhood pixel q .

$$D(p) = \sum_{r \in N_p} dw(r)/4$$

is the displacement confidence term that denotes the reliable displacement information contained in p 's 4-neighborhood N_p , where $dw(r)$ is the displacement weight of p 's 4-neighborhood pixel r .

$$S(p) = \nabla C_p^\perp \cdot n_p / \alpha$$

is the structure term, in which ∇C_p describes the maximum color gradient in NB_p , \perp denotes the orthogonal operator. n_p is the unit normal of p on $\partial\Omega_h$ and α is a normalization factor. $S(p)$ represents the interaction strength of the

image structure with the boundary of the hole, and boosts the priority of a fragment where the structural interaction happens [CPT04].

Table 2 shows the pseudo-code of energy optimization for hole filling. In our experiments, λ_c and λ_e are fixed in the range of $0.5 \sim 1.5$ and $1 \sim 2$. λ_s is initialized within range of $1 \sim 2$ and increased by 1 after each Loop 1 to enforce the constraint of displacement field smoothness. With the algorithm in Table 2, all missing pixels in the hole are repaired with the optimized corresponding pixels on the LDV image.

3.4. Elimination of ghost effect

The missing regions on the target image are faithfully repaired by applying the above distortion minimization. However, due to the luminance difference between the LDV image and the target image, ghost effect may appear on the repaired result as shown in Fig. 6(g). We eliminate this phenomenon by Poisson image blending [PGB03].

Suppose f^* is the known color of all pixels in $\delta\Omega_h$, we obtain the fusion color f for all pixels in Ω_h by solving the following linear equations:

$$|N_p|f_p - \sum_{q \in N_p \cap \Omega_h} f_q = \sum_{q \in N_p \cap \delta\Omega_h} f_q^* + \sum_{q \in N_p \cap \Omega_h} g_{pq},$$

where $p \in \Omega_h$, $g_{pq} = c_p - c_q$ is the color gradient between the repaired pixels p and q in Ω_h . We adopt the bi-conjugate gradient method to solve the above large-scale sparse linear equations with high efficiency.

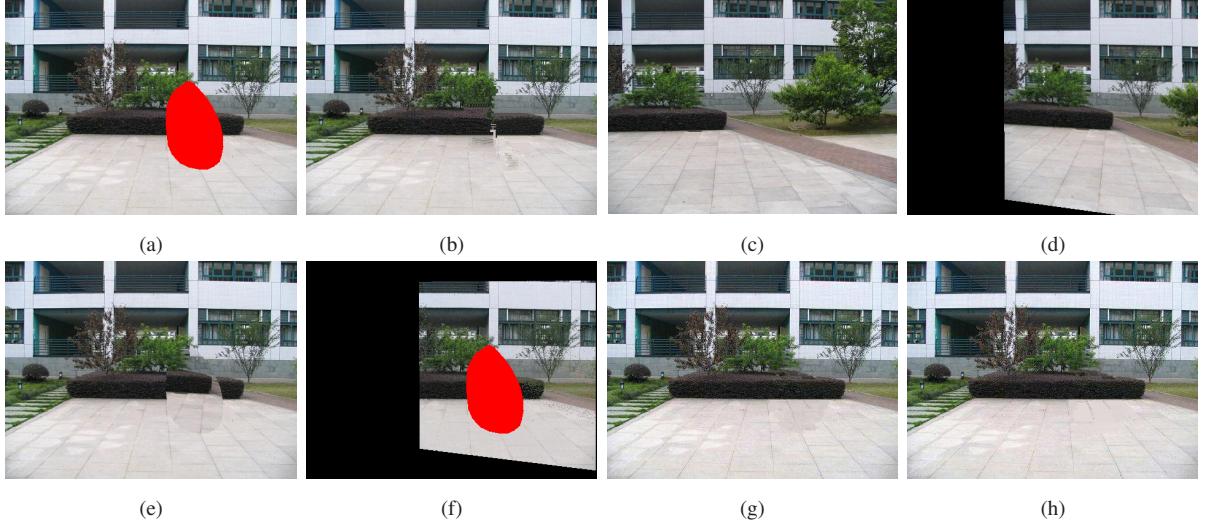


Figure 6: Completion of the damaged region. (a) The target image with the missing pixels in red; (b) The repairing result using a previous method [CPT04]; (c) The LDV image; (d) The warped LDV image; (e) Image stitching result with obvious mismatch on the ground and the shrub; (f) Optimization result of overlap correspondences; (g) The result without Poisson image blending; (h) The final repairing result.

4. Results and discussions

We implemented the proposed algorithm on an Intel Pentium IV 1.8GHz PC with 1GB main memory under the Windows XP operating system. Four experimental results are demonstrated here.

Fig.1 concerns human removal. We first adopt a classical texture synthesis based image completion method [CPT04] to repair the occluded region on the target image Fig.1(a) and get Fig.1(b). Obviously, the tree's structure and some distinct features cannot be well restored due to the illness nature of the method [She03]. Besides, interactive image completion techniques [SLJS05, PSK06] are also unfeasible for this example, as they need to carefully specify structure curves which is an impossible task for such a missing region with complex structures. By introducing a LDV image of the same scene as Fig.1(d), we aim at a more natural repairing result. Nevertheless, the previous approach in [LGP*07] fails for this example due to the difficulties of setting sufficient reliable feature correspondences on the ground.

After applying homography based LDV warping, the warped result is shown in Fig.1(e). Compared with Fig.1(a), distinct distortions still exhibit on the wall and the floor because of their large displacement view change. Repairing the occluded region in Fig.1(c) with Fig.1(e) directly causes poor repairing result [Sze04] in Fig.1(f). Our algorithm further corrects the remaining perspective distortion in the common scene regions between Fig.1(a) and Fig.1(d) by energy optimization. The optimized overlap correspondences is displayed in Fig.1(g), in which the corrected common known scene region on the LDV image is harmonious with that on the target image. Finally, the initial estimate of the miss-

ing region is rectified by energy optimization under the constraints of reliable overlap correspondences. Fig.1(h) is our repairing result, which is the best compared with Fig.1(b) and Fig.1(f). It took less than 1 minute to repair about 9,000 missing pixels on the target image with the size of 461×346 .

In a similar way, Fig.4 and Fig.5 show the removal of obstacles in front of the scene of interest, with 9,284 and 19,727 occluded pixels respectively. They show that our algorithm works well even for the large missing regions with complex structure information.

The last example of repairing a damaged region with 10,170 missing pixels is shown in Fig.6. Obvious ghost effect exists due to the illuminance difference between the target image and the LDV image as shown in Fig.6(g). Fig.6(h) shows the seamless repairing result with Poisson image blending. This example proves that our algorithm can produce good repairing result when there are slight luminance difference between the two views.

More experimental results and their comparisons with the ground truths can be found in the attached multimedia file.

5. Conclusions and future work

This paper presents a new image completion method based on an additional LDV image for faithfully repairing large missing regions on the target image in an automatic way. A coarse-to-fine distortion correction algorithm is proposed to minimize the perspective distortion in the LDV image, which consists of the following steps. First, homography based LDV warping provides an initial distortion correction of the common known scene regions on the LDV image and

Table 2: Energy Optimization for Hole Filling.

-
1. Initialization.
 - Initialize p' with inverse homography H^{-1} , i.e., $p' \sim H^{-1}p, \forall p \in \Omega_h$.
 - Initialize B_h with $\partial\Omega_h$.
 2. Weight setup.
 - If $p_i \in \delta\Omega_h \cup R$, $cw(p_i) = 1$; else $cw(p_i) = 0$.
 - If $p_i \in (\delta\Omega_h \cap M_g) \cup R$, $dw(p_i) = 1$; else $dw(p_i) = 0$.
 3. Priority computation.
 - Compute $C(p), D(p), S(p), \forall p \in B_h$.
 - If $C(p) < \theta_c \parallel D(p) < \theta_d \parallel S(p) < \theta_s$, $P(p) = 0$; else $P(p) = C(p) * D(p) * S(p)$.
 4. Energy function minimization.

For $p_m = \arg \max_{p \in B_h} [P(p) > 0]$, update p'_m by minimizing related energy function $E(p_m)$ using the conjugate gradient (CG) method.
 5. Copying and filling.

Repair p_m with p'_m , set $cw(p_m) = 0, dw(p_m) = 1$.
 6. Update of B_h and R .
 7. Optimization loops.
 - *Loop 1.* Repeat steps 4 ~ 6 until none is repaired in two successive iterations.
 - If $B_h = \emptyset$, exit, else increase λ_s and go to *Loop 2*.
 - *Loop 2.* Repeat steps 2 ~ 6 until $B_h = \emptyset$.
-

an initial estimate of the missing pixels on the target image. Second, the residual distortions in the common known scene regions of the LDV image are further rectified by energy optimization of pixel correspondences in the overlap. Third, under the constraints of epipolar geometry, displacement field smoothness and color consistency of the neighboring pixels, the missing pixels are orderly repaired according to a specially defined priority function. Experiments show that our method work well even for the large missing region with complex structure, and obtains the repairing result superior to previous image completion techniques.

At present, we only verify our method with a single LDV image. However, image completion based on multiple views will be more flexible for photo editing and useful for completion with complex occlusions. This is the future work we will continue to do.

Acknowledgement

This work was supported by National Natural Science Foundation of China under Grant No. 60403038, No. 60703084 and No. 60603076, National 863 High Technology Plan Foundation of China under Grand No. 2007AA01Z316, and

National Grand Foundation Research 973 Program of China under Grand No. 2002CB312101. The authors would like to thank the anonymous reviewers for their professional guidance and detailed comments on this paper.

References

- [AMN*98] ARYA S., MOUNT D.-M., NETANYAHU N.-S., SILVERMAN R., WU A.: An optimal algorithm for approximate nearest neighbor searching. *Journal of ACM* 45, 6 (Nov. 1998), 891–923.
- [BB95] BEAUCHEMIN S.-S., BARRON J.-L.: The computation of optical flow. *ACM Computing Surveys (CSUR)* 27, 3 (Sep. 1995), 433–466.
- [BBS01] BERTALMIO M., BERTOZZI A.-L., SAPIRO G.: Navier-stokes, fluid dynamics, and image and video inpainting. In *Proceedings of the IEEE CVPR 2001* (Dec. 2001), vol. 1, pp. 355–362.
- [BCV*01] BALLESTER C., CASELLES V., VERDERA J., BERTALMIO M., SAPIRO G.: A variational model for filling-in gray level and color images. In *Proceedings of the IEEE ICCV 2001* (Jul. 2001), vol. 1, pp. 10–16.
- [BSCB00] BERTALMIO M., SAPIRO G., CASELLES V., BALLESTER C.: Image inpainting. In *Proceedings of the ACM SIGGRAPH 2000* (Jul. 2000), pp. 417–424.
- [CK04] COLLIS B., KOKARAM A.: Filling in the gaps. *IEE Electronics Systems and Software* 2, 4 (Aug./Sep. 2004), 22–28.
- [CPT04] CRIMINISI A., PÉREZ P., TOYAMA K.: Region filling and object removal by exemplar-based image inpainting. *IEEE Transactions on Image Processing* 13, 9 (Sep. 2004), 1200–1212.
- [CS05] CHAN T., SHEN J.-H.: Variational image inpainting. *Communications on Pure and Applied Mathematics* 58, 5 (Feb. 2005), 579–619.
- [DCOY03] DRORI I., COHEN-OR D., YESHURUM H.: Fragment-based image completion. *ACM Transactions on Graphics* 22, 3 (Jul. 2003), 303–312.
- [FS05] FADILI M.-J., STARCK J.-L.: EM algorithm for sparse representation-based image inpainting. In *Proceedings of the IEEE ICIP 2005* (Sep. 2005), vol. 2, pp. 61–64.
- [GS98] GAO P., SEDERBERG T.-W.: A work minimization approach to image morphing. *The Visual Computer* 14, 8-9 (Dec. 1998), 390–400.
- [HE07] HAYS J., EFROS A.-A.: Scene completion using millions of photographs. *ACM Transactions on Graphics (TOG)* 26, 3 (Jul. 2007), 4–1–4–4.
- [HZ00] HARTLEY R., ZISSERMAN A.: *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049, Canberra, Australia, Jul. 2000.

- [JHM05] JIA Y.-T., HU S.-M., MARTIN R.-R.: Video completion using tracking and fragment merging. *The Visual Computer* 21, 8-10 (Sep. 2005), 601–610.
- [JT03] JIA J.-Y., TANG C.-K.: Image repairing: Robust image synthesis by adaptive ND tensor voting. In *Proceedings of the IEEE CVPR 2003* (Ju. 2003), vol. 1, pp. 643–650.
- [JTW06] JIA J.-Y., TAI Y.-W., WU T.-P., TANG C.-K.: Video repairing under variable illumination using cyclic motions. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28, 5 (May 2006), 832–839.
- [KT06] KOMODAKIS N., TZIRITAS G.: Image completion using global optimization. In *Proceedings of the IEEE CVPR 2006* (Ju. 2006), vol. 1, pp. 442–452.
- [LGP*07] LIU C.-X., GUO Y.-W., PAN L., PENG Q.-S., ZHANG F.-Y.: Image completion based on views of large displacement. *The Visual Computer* 23, 9-11 (Jun. 2007), 833–841.
- [Low04] LOWE D.-G.: Distinctive image features from scale-invariant interest points. *International Journal of Computer Vision* 60, 2 (Nov. 2004), 91–110.
- [LZW03] LEVIN A., ZOMET A., WEISS Y.: Learning how to inpaint from global image statistics. In *Proceedings of the IEEE ICCV 2003* (Oct. 2003), pp. 305–312.
- [OABB85] OGDEN J.-M., ADELSON E.-H., BERGEN J.-R., BURT P.-J.: Pyramid-based computer graphics. *RCA Engineer* 30, 5 (1985), 4–15.
- [PGB03] PÉREZ P., GANGNET M., BLAKE A.: Poisson image editing. *ACM Transactions on Graphics* 22, 3 (Jul. 2003), 313–318.
- [PSB07] PATWARDHAN K.-A., SAPIRO G., BERTALMIO M.: Video inpainting under constrained camera motion. *IEEE Transactions on Image Processing* 16, 2 (Feb. 2007), 545–553.
- [PSK06] PAVIĆ D., SCHÖNEFELD V., KOBELT L.: Interactive image completion with perspective correction. *The Visual Computer* 22, 9-11 (September 2006), 671–681.
- [SC05] SHIH T.-K., CHANG R.-C.: Digital inpainting - survey and multilayer image inpainting algorithms. In *Proceedings of ICITA 2005* (Jul. 2005), vol. 1, pp. 15–24.
- [SD96] SEITZ S.-M., DYER C.-R.: View morphing. In *Proceedings of the ACM SIGGRAPH 1996* (Aug. 1996), pp. 21–30.
- [She03] SHEN J.-H.: Inpainting and the fundamental problem of image processing. *SIAM News* 36, 5 (Ju. 2003), 1–4.
- [SLJS05] SUN J., LU Y., JIA J.-Y., SHUM H.-Y.: Image completion with structure propagation. *ACM Transactions on Graphics* 24, 3 (Jul. 2005), 861–868.
- [SMKT06] SHIRATORI T., MATSUSHITA Y., KANG S.-B., TANG X.: Video completion by motion field transfer. In *Proceedings of the IEEE CVPR 2006* (Ju. 2006), vol. 1, pp. 411–418.
- [SS02] SCHARSTEIN D., SZELISKI R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision* 47, 1-3 (Apr. 2002), 7–42.
- [STG03] STRECHA C., TUYTELAARS T., GOOL L. V.: Dense matching of multiple wide-baseline views. In *Proceedings of the IEEE ICCV 2003* (Oct. 2003), vol. 2, pp. 1194–1201.
- [Sze04] SZELISKI R.: *Image Alignment and Stitching: A Tutorial*. Tech. Rep. MSR-TR-2004-92, Microsoft Research, Microsoft Corporation, One Microsoft Way, Redmond, WA, Dec. 2004.
- [WSI04] WEXLER Y., SHECHTMAN E., IRANI M.: Space-time video completion. In *Proceedings of the IEEE CVPR 2004* (Ju. 2004), vol. 1, pp. 120–127.