Chunxiao Liu · Yingzhen Yang · Qunsheng Peng · Yanwen Guo

A New Distortion Minimization Approach for Image Completion based on a Large Displacement View

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 LDV image, then the rectified LDV image is used to restore the missing pixels. First, under the assumption of a planar scene, the LDV image is warped according to a homography to generate the initial correction result. Second, by means of a mismatch detection mechanism, the remaining distortion in the common scene regions is relaxed by energy optimization of overlapping correspondences, with the expectations of color constancy and displacement field smoothness. The fundamental matrix for the two views is then computed based on the updated 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 repaired according to a specially defined 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.

Keywords Image completion · Large displacement view · Distortion minimization · Energy optimization · Pixel correspondences

Chunxiao Liu · Yingzhen Yang · Qunsheng Peng
State Key Lab of CAD&CG, Zhejiang University, 310027, Hangzhou, China
Tel.: +86-571-88206680
Fax: +86-571-88206680
E-mail: {liuchunxiao,yangyingzhen,peng}@cad.zju.edu.cn

Yanwen Guo
National Laboratory for Novel Software Technology, Nanjing University, 210093, Nanjing, China
Tel.: +86-25-83592398
E-mail: ywguo@graphics.nju.edu.cn

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, rig removal and special effects production [1], image completion has received great attention in the past decade, and many methods have been brought forward so far [2].

Traditional image completion methods can be roughly classified into *PDEs* and 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 regions, but may fail for those with complex texture. The latter methods can produce compelling completion results for relatively large missing region with distinct texture characteristics 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 [3]. 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. We propose here a new image completion method based on an additional large displacement view (LDV) image of the same scene for repairing large missing regions, which gets rid of the above two assumptions.

Although few methods address image completion based on views of large displacement, it provides a feasible way

ISBN: 978-9944-5685-4-8

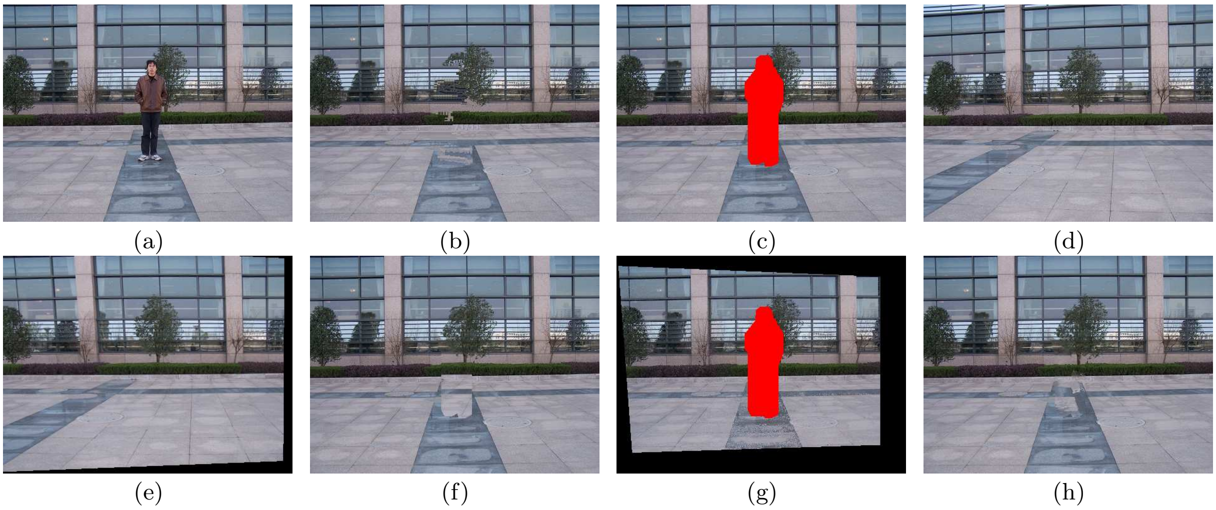

Fig. 1 Completion for the profile of a building. (a) The target image; (b) The repairing result by texture synthesis based method [13]; (c) The occluded region (red) after removing the human; (d) The LDV image; (e) The warped LDV image; (f) Image stitching result [30] with obvious mismatch on the ground; (g) Optimization result of overlapping correspondences; (h) Our repairing result.

for image completion in our life. For instance, when we take photos in some famous resorts, an undesired person may run into the view of our camera. The problem is therefore how to generate the visually pleasing result after removal of the intruder from the image. 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, faithfully enabling a restoration of the occluded regions on the target image.

Basically, we adopt the LDV image to complete the missing regions in the target image. However, it is a non-trivial task to achieve a faithful completion. Two key challenges have to be investigated. The first challenge is how to correct the perspective distortion in 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 with different scene depths. We then need to find better pixel correspondences over the common scene regions of the two views rather than that by global warping. Since traditional stereo matching algorithms [4] and optical flow algorithms [5] only fit for the pixel correspondence problem with short baseline, we propose a new energy optimization scheme for dense pixel correspondences of the LDV views under the constraints of color constancy and displacement field smoothness. During the course of optimization, we adopt a mismatch detection mechanism to reject the lost correspondences and a dynamic increasing weighting parameter to correct them. Considering numerical stability, pixel-wise optimization is adopted.

After distortion correction for the common known scene regions on the LDV image, another challenge is how to exploit the rectified surrounding known regions to estimate the missing regions on the target image. We also treat this problem as energy optimization of pixel correspondences. According to epipolar geometry, the corresponding pixel on the LDV image lies on the epipolar line determined by the current pixel on the target image and a fundamental matrix [6]. 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 repairing 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 [7] to generate seamless result.

Our 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 for the dense pixel correspondences on the LDV images, which is regarded as the correction of remaining distortions in the common known scene regions. Third, under the constraints of epipolar geometry, displacement 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.

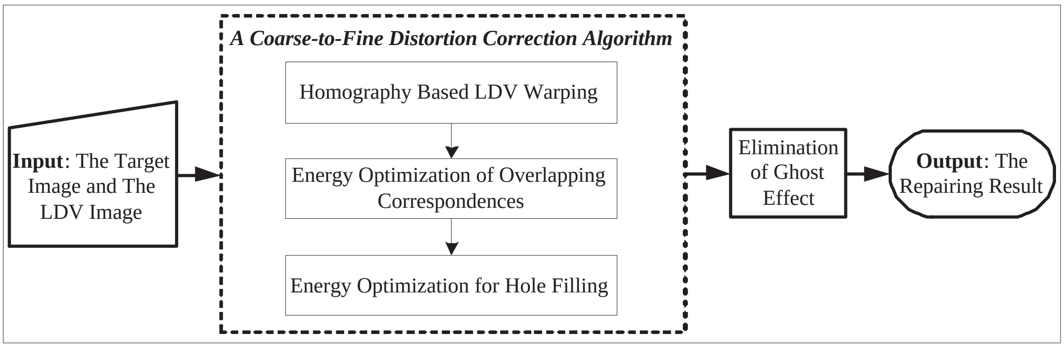

Fig. 2 The algorithm overview.

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.* [8] 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 [8] or variational problems [9,10] 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 regions, 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 [11–14]. 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 [15] or specified projective transformation in the scene [16], remarkable 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 [17] 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 repairing effect.

Statistics based methods solve the problem of image completion by statistical analysis. Levin [18] 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 [19] 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, our previous work [20] for this problem is limited for the quasi-planar scenes and needs the tiresome effort for the interactive segmentation of planar scene regions.

2.2 Video completion

Bertalmio *et al.* [21] extended PDEs based image completion to video sequences. 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 [22] and [23] recover the missing information by direct sampling spatio-temporal patches of local structures or motion. Jia *et al.*'s method [24] 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 [25]. The recent work by Patwardhan *et al.* [26] first segments each frame into static background and moving foreground, then perform texture synthesis based motion 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 adjacent frames for repairing. However, for the problem discussed in this paper, we cannot directly use the information on the LDV image due to the severe distortion of the same scene in the two views.

3 Distortion minimization for image completion based on a LDV image

In this paper, we propose an algorithm for repairing the large missing regions on an image based on distortion minimization of another view with large displacement. 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.

A coarse-to-fine distortion correction algorithm is proposed to correct perspective distortion in the common known scene regions on the LDV image, and to estimate the missing regions on the target image. 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

We first assume that the scene in the two views is approximatively located on a 3D plane. The LDV image can therefore be globally warped through a homography matrix to roughly estimate their common scene regions. The warped LDV image establishes an initial distortion correction for the common known scene regions on the two views. It also provides an initial estimate for the missing pixels on the target image. We will further rectify them in the following subsections.

During homography based LDV warping, our approach proceeds by the following steps in order:

(1)*Feature Detection*: A robust feature point extraction method, i.e. scale invariant feature transform (*SIFT*) feature detector [27], is employed to obtain enough feature points and their high-dimensional feature descriptors on the two views.

(2)*Feature Matching*: The approximate nearest neighbor (ANN) searching algorithm [28] is exploited to find 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 the influence of image noise. We adopt the RANSAC algorithm and the Levenberg-Marquardt algorithm [6] 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, $h_0, ..., h_7$ 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.

The warped LDV image S' falls on the target image T. The part of S' which overlaps with the known regions on the target image establishes an initial pixel correspondence for the common known scene region in the overlappings Ω_o. As S' covers the missing region Ω_h, it in fact provides an initial estimate of the missing pixels.

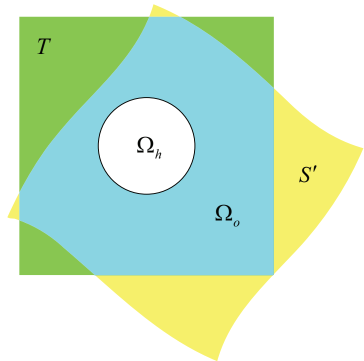

Fig. 3 Homography based LDV warping. The warped LDV image S' is pasted onto the target image T, providing the initial estimate for the missing pixels in the hole Ω_h and the initial distortion correction for the common known scene region in the overlappings Ω_o.

3.2 Energy optimization of overlapping correspondences

Since the warping is based on the assumption that the scene geometry is nearly planar, there are a large number of mismatches in the overlappings Ω_o. According to epipolar geometry [6], it would be more reasonable to take the fundamental matrix to depict the geometrical relationship between two views. To accomplish this, two problems need to be resolved. The first problem is how to compute the fundamental matrix based on pixel correspondences in the overlappings while avoiding the gross errors caused by mismatches. Note that the fundamental matrix is not a sufficient displacement constraint for estimating the missing pixels. We regard the displacement field smoothness as an additional constraint. So the second problem is how to estimate the displacement field with the existence of mismatches in the overlappings.

Solving the above two problems necessitates further correction of the remaining distortions in the overlappings. Traditional stereo matching algorithms [4] and optical flow algorithms [5] are only suitable for the dense correspondence problem with short baseline. With the help of a mismatch detection mechanism and a dynamic weighting parameter, we formulate distortion correction in the overlappings as a problem of energy optimization for pixel correspondences and 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 on the LDV image p'. Let N_p represent p's 4-connected neighbors in Ω_o, and $< p, q >$ is a pixel pair such that $q \in N_p$. We model the overlapping 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

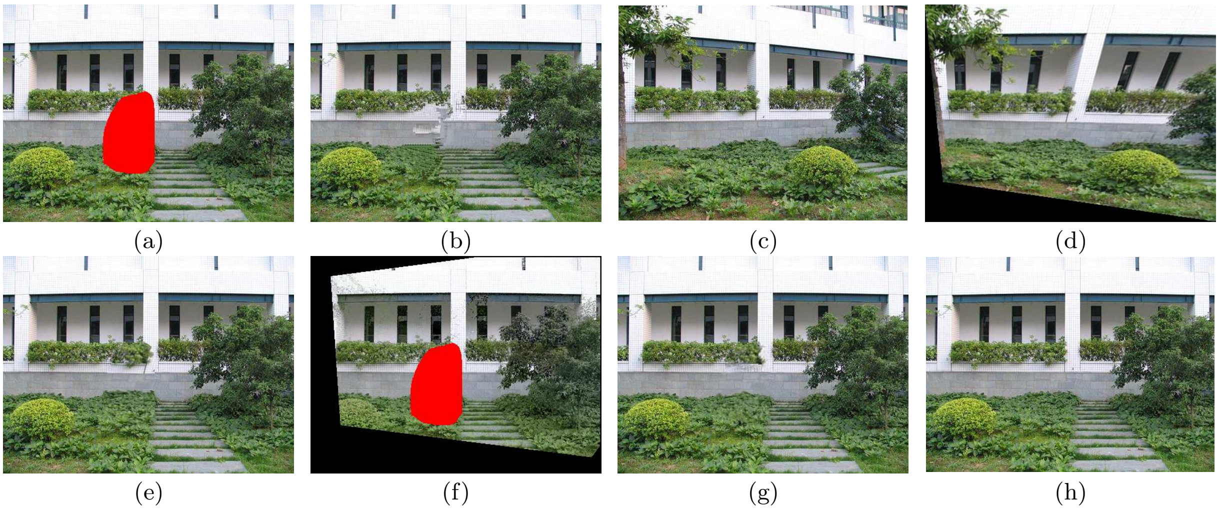

Fig. 4 Completion for one corner of a building. (a) The target image with a red missing region; (b) Repairing result by texture synthesis based method [13]; (c) The LDV image; (d) The warped LDV image; (e) Image stitching result with slight mismatch on the wall; (f) Optimization result of overlapping correspondences; (g) Our repairing result; (h) The ground truth.

displacement field smoothness within neighboring pixels, we define the energy function as

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

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 $E_s(p,q)$.

In Eq.(1), the color constancy term ensures that the corresponding pixel color $C'_{p'}$ on the LDV image agrees with C_p on the target image. As a result, 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 [5] between the corresponding pixel coordinates p' and q', i.e.,

$$E_s(p,q) = (p' - q')^2.$$

However, it 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 $\rho(x)$. Our displacement field smoothness term can thus be expressed as

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

thereinto

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

For the irregularity of the image data and the large scale unknowns, simultaneous optimization of displacement vectors for all the overlappings not only often traps in the unstable solution with local minimum but also is usually computationally intensive. Inspired by the work in [29], a relaxed optimization strategy is adopted here to solve this problem, i.e., only optimizing the displacement vector of one pixel at one time and the others are fixed. Considering the energy function Eq.(1), we can see that only several 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

$$\begin{aligned} E(p) = (C'_{p'} - C_p)^2 + \lambda[\rho(X_{i,j} - X_{i-1,j} - 1) + \\ \rho(X_{i,j} - X_{i,j-1}) + \rho(X_{i+1,j} - X_{i,j} - 1) + \\ \rho(X_{i,j+1} - X_{i,j}) + \rho(Y_{i,j} - Y_{i-1,j}) + \\ \rho(Y_{i,j} - Y_{i,j-1} - 1) + \rho(Y_{i+1,j} - Y_{i,j}) + \\ \rho(Y_{i,j+1} - Y_{i,j} - 1)]. \end{aligned} \quad (4)$$

During 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 overlapping correspondences is summarized in Table.1.

In Table.1, a strict mismatch detection mechanism is devised to reject those lost 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. $ColorDiffer(p)$ means the absolute color difference between p on the target image and p' on the LDV image. During optimiza-

Table 1 Energy Optimization of Overlapping 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 $DisplaceTerm(p) > \theta_{dt}$ &&
   $(ColorSimilarity(p) > \theta_{cs}$ ||
   $ColorDiffer(p) > \theta_{cd})$,
   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 $3 \sim 4$.
   - If p violates 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.

tion, 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 correspondences.

With enough reliable pixel correspondences $(p, p') \in M_g$ in the overlappings, we can estimate a fine fundamental matrix F for the two views via the normalized 8-point algorithm [6]. 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 via $p'^{\mathbf{T}}Fp = 0$, i.e.,

$$\begin{pmatrix} X_{i,j} & Y_{i,j} & 1 \end{pmatrix} \cdot \begin{pmatrix} f_0 & f_1 & f_2 \\ f_3 & f_4 & f_5 \\ f_6 & f_7 & 1 \end{pmatrix} \cdot \begin{pmatrix} i \\ j \\ 1 \end{pmatrix} = 0,$$

where $f_0, ..., f_7$ 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 a reliable correspondence set M_g in the overlappings, 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, 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., $p'^{\mathbf{T}}Fp = 0$. 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 overlappings. 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.,

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

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 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 pixel-wise, and begins with the boundary pixels of the hole B_h with high repairing priority. For the sake of reliable hole filling and structure preserving, the repairing priority for $p \in B_h$ is defined 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.

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

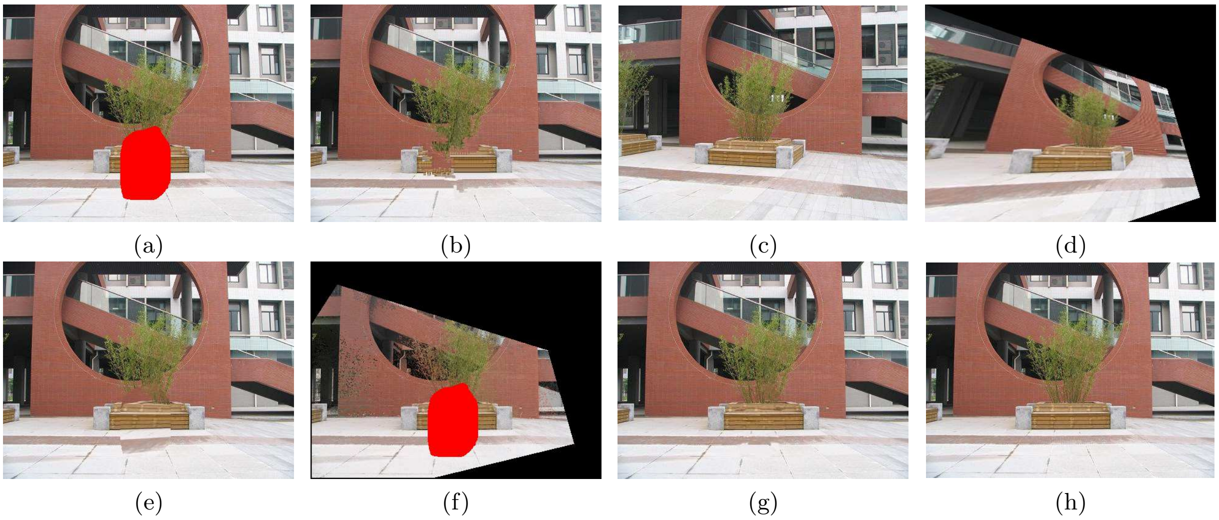

Fig. 5 Completion for one corner of a building. (a) The target image with the damaged region in red; (b) The repairing result using a classical method [13]; (c) The LDV image; (d) The warped LDV image; (e) Image stitching result with obvious mismatch on the ground and incompleteness in the bamboo; (f) Optimization result of overlapping correspondences; (g) Our repairing result; (h) The ground truth.

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 [13].

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 $C(p) < \theta_c \parallel D(p) < \theta_d \parallel Sp < \theta_s$, $P(p) = 0$; 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 $\sim$ 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 $\sim$ 6 until $B_h = \emptyset$.

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 enhance 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 with the above distortion minimization process. 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 [7].

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.

4 Results and discussions

We implemented the proposed algorithm on an Intel Pentium IV 1.8GHz PC with 1GB main memory under the

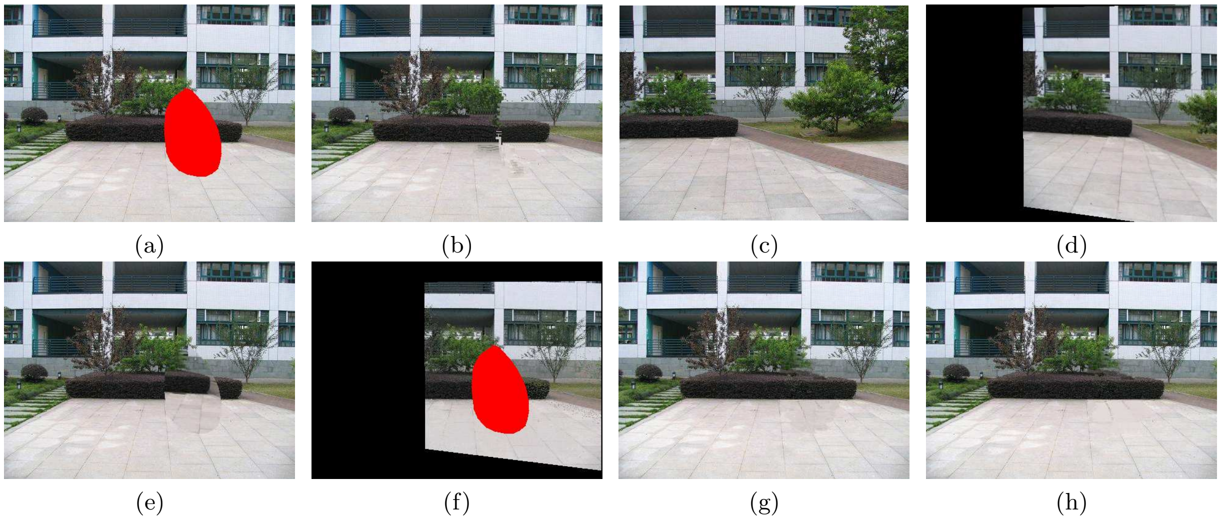

Fig. 6 Completion for one corner of a building. (a) The target image with red missing region; (b) The repairing result using a previous method [13]; (c) The LDV image; (d) The warped LDV image; (e) Image stitching result with obvious mismatch on the ground and the shrub; (f) Optimization result of overlapping correspondences; (g) The result without Poisson image blending; (h) The final repairing result.

Windows XP operating system. Four experimental results are demonstrated here.

For object removal in Fig.1, we first adopt a classical texture synthesis based image completion method [13] 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 [3]. In addition, interactive image completion techniques [15,16] are unfeasible for this example, as they need to carefully specify structure curves which is impossible 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 [20] fails for this example due to lack of feature correspondences on the ground.

Through homography based LDV warping, the warped LDV image 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 [30] 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 overlapping correspondences is displayed in Fig.1(g), in which the corrected overlapping region on the LDV image is harmonious with that of the target image. Finally, the initial estimate of the missing region is rectified by energy optimization under the constraints of good overlapping correspondences. Fig.1(h) is our repairing result, which is the best one 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.

Fig.4 and Fig.5 demonstrate two repairing examples for damaged images with 9,822 and 8,863 missing pixels respectively. For each of them, (a) shows the target image with the missing region in red. (b) presents the poor repairing result by traditional texture synthesis based methods [13]. (c) is the corresponding LDV image. (d) is the warped LDV image with remaining distortions, therefore disharmonious repairing result is generated in (e) through image stitching techniques. As a result, further optimization of the remaining distortions in the overlappings is necessary and the correction result is shown in (f). With energy optimization, our repairing result is shown in (g), which is comparable to the ground truth in (h). As can be seen, our algorithm works well even for the large missing regions with complex structure information.

The last experimental result with 10,170 missing pixels is shown in Fig.6. Fig.6(g) shows the repairing result without Poisson image blending. Obvious ghost effect exists due to the illuminance difference between the target image and the LDV image. In comparison, 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.

Please see more detailed experimental results in the attached 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, then the rectified LDV image is used to restore the missing pixels. First, homography based LDV warping provides an initial distortion correction of the common known scene regions and an initial estimate of the missing pixels. Second, the remaining distortion in the common known scene regions of the LDV image is further rectified by energy optimization of pixel correspondences. 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 achieves 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 and useful for photo editing and completion. We'll further test the cases of several LDV images with complex occlusions. This is the future work we concern.

Acknowledgement

This work was supported by National Natural Science Foundation of China under Grant No. 60403038, National 863 High Technology Plan Foundation of China under Grand No. 2006AA01Z314, and National Grand Foundation Research 973 Program of China under Grand No. 2002CB312101. The authors benefit from the stimulating discussions with Dr. Guofeng Zhang.

References

1. Collis, B., Kokaram, A.: Filling in the Gaps. IEE Electronics Systems and Software. 2(4), 22-28 (2004)
2. Shih, T.K., Chang, R.C.: Digital Inpainting - Survey and Multilayer Image Inpainting Algorithms. In: Proceedings of ICITA 2005, pp. 15-24 (2005)
3. Shen, J.H.: Inpainting and the Fundamental Problem of Image Processing. SIAM News. 36(5), 1-4 (2003)
4. Scharstein, D., Szeliski, R.: A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. International Journal of Computer Vision. 47(1-3), 7-42 (2002)
5. Beauchemin, S.S., Barron, J.L.: The Computation of Optical Flow. ACM Computing Surveys (CSUR). 27(3), 433-466 (1995)
6. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press, ISBN: 0521623049. (2000)
7. Pérez, P., Gangnet, M., Blake, A.: Poisson Image Editing. ACM Transactions on Graphics. 22(3), 313-318 (2003)
8. Bertalmio, M., Sapiro, G., Caselles, V., Ballester, C.: Image Inpainting. In: Proceedings of ACM SIGGRAPH 2000, pp. 417-424. New Orleans, Louisiana (2000)
9. 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. 1, 10-16 (2001)
10. Chan, T., Shen, J.H.: Variational Image Inpainting. Communications on Pure and Applied Mathematics. 58(5), 579-619 (2005)
11. Drori, I., Cohen-Or, D., Yeshurum, H.: Fragment-Based Image Completion. ACM Transactions on Graphics. 22(3), 303-312 (2003)
12. Jia, J.Y., Tang, C.K.: Image Repairing: Robust Image Synthesis by Adaptive ND Tensor Voting. In: Proceedings of the IEEE CVPR 2003. 1, 643-650 (2003)
13. Criminisi, A., Pérez, P., Toyama, K.: Region Filling and Object Removal by Exemplar-Based Image Inpainting. IEEE Transactions on Image Processing. 13(9), 1200-1212 (2004)
14. Komodakis, N., Tziritas, G.: Image Completion Using Efficient Belief Propagation Via Priority Scheduling and Dynamic Pruning. IEEE Transactions on Image Processing. 16(11), 2649-2661 (2007)
15. Sun, J., Lu, Y., Jia, J.Y., Shum, H.Y.: Image Completion with Structure Propagation. ACM Transactions on Graphics. 24(3), 861-868 (2005)
16. Pavić, D., Schönefeld, V., Kobbelt, L.: Interactive image completion with perspective correction. The Visual Computer (PG 2006). 22(9-11), 671-681 (2006)
17. Hays, J., Efros, A.A.: Scene Completion Using Millions of Photographs. ACM Transactions on Graphics (TOG). 26(3), 4-1-4-4 (2007)
18. Levin, A., Zomet, A., Weiss, Y.: Learning How to Inpaint from Global Image Statistics. In: Proceedings of the IEEE ICCV 2003, pp.305-312. Nice, France (2003)
19. Fadili, M.J., Starck, J.L.: EM Algorithm for Sparse Representation-Based Image Inpainting. In: Proceedings of the IEEE ICIP 2005. 2, 61-64 (2005)
20. 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), 833-841 (2007)
21. Bertalmio, M., Bertozzi A.L., Sapiro, G.: Navier-Stokes, Fluid Dynamics, and Image and Video Inpainting. In: Proceedings of the IEEE CVPR 2001. I, 355-362 (2001)
22. Wexler, Y., Shechtman, E., Irani, M.: Space-Time Completion of Video. IEEE Transactions on Pattern Analysis and Machine Intelligence. 29(3), 463-476 (2007)
23. Shiratori, T., Matsushita, Y., Kang, S.B., Tang, X.: Video Completion by Motion Field Transfer. In: Proceedings of the IEEE CVPR 2006. 1, 411-418 (2006)
24. Jia, Y.T., Hu, S.M., Martin, R.R.: Video Completion Using Tracking and Fragment Merging. The Visual Computer (PG 2005). 21(8-10), 601-610 (2005)
25. Jia, J., 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), 832-839 (2006)
26. Patwardhan, K.A., Sapiro, G., Bertalmio, M.: Video Inpainting Under Constrained Camera Motion. IEEE Transactions on Image Processing. 16(2), 545-553 (2007)
27. Lowe, D.G.: Distinctive Image Features from Scale-Invariant Interest Points. International Journal of Computer Vision. 60(2), 91-110 (2004)
28. 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), 891-923 (1998)
29. Gao, P., Sederberg, T.W.: A Work Minimization Approach to Image Morphing. The Visual Computer. 14(8-9), 390-400 (1998)
30. Szeliski, R.: Image Alignment and Stitching: A Tutorial. In: Microsoft Research Technical Reports, http://research.microsoft.com/vision/visionbasedmodeling/publications/MSR-TR-2004-92-Jan26.pdf. (2004)