Image Completion Using Structural Priority Belief Propagation

Yingzhen Yang
State Key Lab of CAD&CG
Zhejiang University
Hangzhou, China
yangyingzhen@cad.zju.edu.cn

Yin Zhu
National Key Laboratory for
Novel Software Technology
Nanjing University
Nanjing, China
zhuyin.nju@gmail.com

Qunsheng Peng
State Key Lab of CAD&CG
Zhejiang University
Hangzhou, China
peng@cad.zju.edu.cn

ABSTRACT

A new image completion algorithm called Structural Priority Belief Propagation (SPBP) is presented to deal with LDV based image completion in this paper. LDV completion is a new form of image completion based on another large displacement view (LDV) of the same scene, no wonder, it has the potential of repairing large unknown region with salient structure information. In order to complete such unknown region, SPBP makes two important extensions over existing Priority-BP: dynamic weight of structural consistency and structural priority inheritance so as to propagate linear structure with correct priority, meanwhile it promotes texture propagation adhering to a global optimization scheme. Experimental results demonstrate that SPBP can obtain more satisfactory results than other LDV completion algorithms and it also performs well for traditional single image completion.

Categories and Subject Descriptors: I.4 [Image Processing and Computer Vision]: Miscellaneous

General Terms: Algorithm, Experimentation

1. INTRODUCTION

Traditional approaches of single image completion fill in the unknown regions in an image by exploiting the information in the same image to achieve a visually plausible effect. Actually in many cases it is impossible to find proper pixels from the same image to fill in the unknown region. Recently a new form of image completion based on a large displacement view of the same scene (LDV image) is proposed in [10] and [11]. Given a target image and a LDV image, the objective of LDV completion is to repair the large unknown region Ω in target image which may contain salient structure information provided that the unknown region Ω is visible in LDV image. Fig. 1 is an example of LDV completion (the man is removed and the unknown region behind him is restored). Since our method of LDV completion is closely related to techniques of single image completion, we briefly review single image completion techniques first.

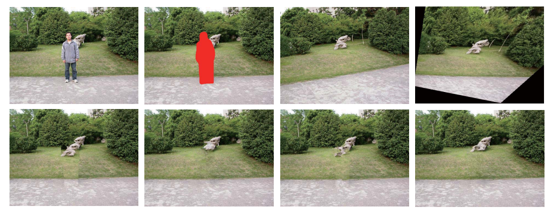

Figure 1: An example of LDV completion. The first row, from left to right: original target image, masked target image, LDV image, and warped LDV image. The second row, from left to right are four completion results: using warped LDV image, Criminisi's method, Liu's method, and our method.

Among various single image completion approaches, PDE based methods in [2] can complete small scratches but they cause blurring in large unknown region. Exemplar-based techniques prove to be more effective by copying source patches or pixels from the known region to fill the unknown region. Jia et al. [8] used a tensor voting algorithm to link the structure in the unknown region. In [4], an iterative method is proposed to complete unknown region using adaptive example fragments, but it is time consuming. Later, researchers realized the importance of completion order for effectively propagating structure information. In [13], the structure propagation which is guided by user specified curves is performed before texture propagation, but the user interaction is tiresome. Criminisi et al. [3] proposed a priority order to encourage propagation of linear structure, but this greedy scheme causes visual inconsistencies as mentioned in [9] and [13]. Contrary to the greedy scheme, global optimization scheme is proposed in [9] and [7]. The Priority-BP in [9] accounts for no structure information thus may cause structural inconsistency, and the structural consistency cost in [7] is simply the difference of gradient which is insufficient to represent high level structural consistency of edges and ridges, and it ignores the completion order. Recent data-driven work in [6] requires a huge number of photos which is difficult for personal use, and the method in [1] needs tiresome user interaction.

[10] and [11] address LDV completion. [10] requires tiresome interactive segmentation and it assumes that the unknown region is quasi-planar. [11] removes this limitation, but it has two drawbacks: first, it performs optimization for each unknown pixel thus is computationally intensive

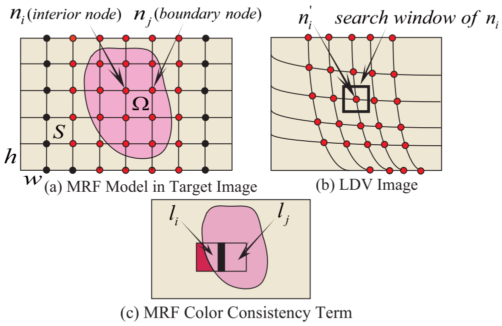

Figure 2: (a) Illustration of MRF Model. (b) The point n_i' in LDV image is corresponding to node n_i by a homography. (3) The color consistency term for label l_i is an SSD in the red region while the same term for label pair (l_i, l_j) is an SSD in the black overlap region between the two labels.

when completing a large unknown region. Second, it uses the greedy scheme in [3] which may introduce visual inconsistency. It is also worthwhile to mention that traditional stereo matching algorithms are only suitable for dense pixel correspondences problem with short baseline therefore cannot help LDV completion, while wide-baseline stereo matching method [12] requires multiple views rather than two views in LDV completion and it adopts the computationally intensive structure-from-motion algorithm.

Completion of a large unknown region with salient structure information must be implemented in some priority order as mentioned in [13] . Existing global optimization scheme considers no structure related priority order. Unfortunately, the greedy scheme with priority order may cause visual inconsistencies. SPBP presented in this paper combines the best of two sides: priority order for structure propagation and global optimization scheme based on Priority-BP in [9], without user interaction in [13]. SPBP incorporates priority order into the global optimization scheme through a well defined "structural priority". Two extensions are made over Priority-BP, "dynamic weight of structural consistency" and "structural priority inheritance", which are demonstrated effective in propagating structure with a correct priority while texture is also propagated in the global optimization scheme. Experimental results demonstrate that SPBP can render more satisfactory results than other related algorithms for LDV completion and it also performs well for single image completion.

2. PROBLEM FORMULATION

2.1 MRF Model

To adopt a global optimization scheme with a well defined energy function, we first construct a MRF model. A lattice is attached to the target image with a horizontal and vertical spacing of w and h respectively. The source region S is the area of target image except the unknown region Ω. A grid node whose $2w \times 2h$ neighborhood intersects the unknown region Ω is a MRF node, and the edges form a 4-neighborhood system on the MRF nodes. So we get a undirected graph $G = (V, E)$, with $V = \{n_i\}_{i=1}^N$ indicating all the N MRF nodes and E indicating the edges of the MRF model (see Fig. 2(a)). We use $B(G)$ to denote the set of boundary nodes of G (the nodes whose four neighbors are not all in Ω). Homography solving in [11] is adopted to compute an approximate corresponding point n_i' in LDV image for each node n_i, and the set of labels $L(n_i)$ of node n_i are all the $2w \times 2h$ patches within a predefined search window $W(n_i)$ centered at n_i' (see Fig. 2(b)). For traditional single image completion, the label set for each node is simply all the $2w \times 2h$ patches in source region S. Our task is to infer an optimal labeling $\bar{L} = \{\bar{l}_i\}_{i=1}^N$ ($\bar{l}_i \in L(n_i)$) to minimizes an energy function:

$$E(L) = \sum_{i=1}^{N} E_i(l_i) + \sum_{(n_i,n_j)\in E} E_{ij}(l_i, l_j), \quad L = \{l_i\}_{i=1}^N, \quad (1)$$

where $E_i(l_i)$ is the single node potential for label l_i and $E_{ij}(l_i, l_j)$ is the pairwise potential for (l_i, l_j) (will be defined later). Obviously our MRF model contains loops, and loopy belief propagation can produce satisfactory results for many computer vision applications including image completion. According to [5], max-product BP can be used to minimize the energy function like (1) by iteratively computing messages along each edge in graph G in the following way:

$$m_{ij}^t(l_j) = \min_{l_i \in L(N_i)} (E_i(l_i) + E_{ij}(l_i, l_j) + \sum_{k\in N(i)\backslash j} m_{ki}^{t-1}(l_i)) \quad (2)$$

where $m_{ij}^t(l_j)$ is the message sent from node n_i to node n_j at iteration t, which means how likely n_i believes n_j should choose label l_j. $N(i)\backslash j$ denotes the neighbors of n_i excluding n_j. After T iterations, a belief value is given to each label l_i of node n_i indicating how likely node n_i should choose label l_i:

$$b_i(l_i) = -E_i(l_i) - \sum_{j\in N(i)} mji^T(l_i) \quad (3)$$

The label $\bar{l}_i$ which maximizes $b_i(l_i)$ is chosen as the optimal label for node n_i. Suppose all the label sets $\{L(n_i)\}_{i=1}^N$ have the same size of L, then the running time of T iterations of BP is $O(TNL^2)$, which is prohibitively high if L is large.

2.2 Introduction to SPBP

In [9], priority-based message scheduling requires that the node with highest priority should be the first one to transmit outgoing messages to its neighbors, and we call this priority "message transmission priority" (MTP). Also, the labels of a node with a belief less than a certain threshold are discarded during the dynamic label pruning in [9], and we call this procedure "single label pruning". The MTP of each node in [9] is defined as the reciprocal of the number of its active labels(those labels which remain after pruning).

Actually priority-based message scheduling has potential advantage to propagate structure with a correct priority. Suppose a node n whose active label set $L_s(n)$ has rich structure information owns the highest MTP to transmit messages, then some of its neighbors, say node n', may immediately get "advice" from node n to choose those labels which structurally match $L_s(n)$ as its active labels. If node n' also has a high MTP, then the neighbors of node n' can get similar "advice" from node n' as early as possible, and we can expect similar situation happens repeatedly to propagate structure propagation from node n to node n' and so on. Note that this type of structure propagation is incorporated into the message passing of BP, so the visual inconsistency caused by greedy method is avoided.

The two extensions to existing Priority-BP make above supposition a reality. Actually "dynamic weight of structural consistency" encourages node n' to prefer labels which structurally match active labels of node n, and "structural priority inheritance" guarantees that node n' can also have a high MTP due to its high structural priority "inherited" from node n. We give following definitions beforehand. $SP_n(l)$ is the structure priority of label l of node n, which is a real value between $[0, 1]$ and quantitatively represents the structural information contained in l. $l_{optimal}(n)$ is the optimal label of node n. SP_n is the structural priority of node n, which is defined to be $SP_n(l_{optimal}(n))$ because the optimal label is the most likely choice of node n.

2.3 Dynamic Weight of Structural Consistency

The pairwise consistency term $E_{ij}(l_i, l_j)$ in (1) includes both color consistency term $C(l_i, l_j)$ and structural consistency term $S(l_i, l_j)$ to propagate both structure and texture. $C(l_i, l_j)$ is the sum of squared difference (SSD) of pixel colors in the overlap region between l_i and l_j (see Fig. 1 (c)). Unlike the metric in [7] considering only the difference of gradient, we adopt curvilinear feature matching metric in [14] as our structural consistency term which accounts for high level structure information of edges and ridges. In order to further encourage structure propagation, we propose dynamic weight of structural consistency term based on structural priority. $E_{ij}(l_i, l_j)$, $E_i(l_i)$ in (1) are defined as below:

$$E_{ij}(l_i, l_j) = (1 - \rho(SP_{ij}))C(l_i, l_j) + \rho(SP_{ij})S(l_i, l_j) \quad (4)$$

$$SP_{ij} = \max\{SP_i(l_i), SP_j(l_j)\}$$

$$\rho(x) = \begin{cases} x & x \in [0, \lambda], 0 < \lambda < 1 \\ \lambda & x \in (\lambda, 1] \end{cases}$$

$$E_i(l_i) = (1 - SP(l_{0i}))C(l_i) + SP(l_{0i})S(l_i, l_{0i}) \quad (5)$$

where $C(l_i)$ is the color consistency term for l_i, which is an SSD of pixel color in the overlapping part between l_i and source region S (see Fig. 1(c)). $SP(l_{0i})$ is the structural priority of l_{0i} (see Section 2.4). We use $\lambda = 0.8$ to prevent ignorance of color. Obviously structural similarity is a dominant factor in (4) if at least one of the two labels has high structural priority. According to message update rule (2), the messages sent from a node n prefer labels with structural matching, if labels of node n are of high structural priority. This property is used to show the effectiveness of SPBP in Section 2.5.

2.4 Structural Priority Inheritance

We adopt the priority order in [3] as the structural priority of the boundary patches since this priority is determined accounting for both confidence map and image edges to propagate linear structure. However, this priority is defined in a greedy scheme, thus unsuitable for defining the structural priority of interior nodes in our global optimization scheme. Because the only way to determine the optimal label of an interior node is to receive messages sent from its neighbors, we develop a new mechanism called structural priority inheritance which is synchronous with message update during belief propagation.

The structural priority of a label $l \in W(n)$ of node n is initialized as below:

$$SP_n(l) = \begin{cases} SP(l_0) - S(l, l_0) & n \in B(G) \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where l_0 denotes the boundary patch located at node n and $SP(l_0)$ is the structural priority of l_0. After the message $m_{ij}^t(l_j)$ from node n_i reaches node n_j , $SP_{n_j}(l_j)$ will be updated as below:

$$SP_{n_j}(l_j) = SP_{n_i} - S(l_{optimal}(n_i), l_j) \quad (7)$$

if the right hand side of (7) is larger than the old value of $SP_{n_j}(l_j)$. Note that node n_j chooses a best inheritance source among its neighbors to obtain maximal structure priority of label l_j.

2.5 Neighbor Label Pruning and Structural Priority Based MTP

To accelerate LDV completion we develop a new label pruning mechanism called neighbor label pruning: because the optimal label of a neighbor node n' of node n cannot be far away from $l_{optimal}(n)$ due to spatial continuity, it is safe to discard labels of node n' whose Euclidean distance to $l_{optimal}(n)$ exceed a predefined threshold D before node n' receives message from node n (the node with highest MTP currently). Since a node with high structure priority should also have high MTP to encourage structure propagation, we define the MTP for node n as below:

$$MTP(n) = SP_n + \frac{1}{|A(n)|} \quad (8)$$

where $A(n)$ is the set of active labels of node n after label pruning. Note that a node with higher structural priority, or less active labels indicating that the node is more confident about its labels, has a higher MTP. The pseudocode of SPBP for LDV completion is shown in Table 1.

Suppose a boundary patch l_0 located at the boundary node n_i have a high structural priority $SP(l_0)$. The belief of a active label l_i of n_i after initialization of SPBP in Tabel 1 is

$$b_i(l_i) = -(1 - SP(l_0))C(l_i, l_0) - SP(l_0)S(l_i, l_0) \quad (9)$$

Due to the high weight $SP(l_0)$, $S(l_i, l_0)$ should be small enough to make l_i active after label pruning (with a high belief), thus the structural priority of l_i is also high. Then node n_i possesses a high structural priority and it also possesses a high MTP to send messages according to (8). If a neighbor of node n_i, e.g. node n_j, has some labels, namely $L_s(n_j)$, which structurally match the active labels of n_i, the

Table 1: SPBP for LDV Completion

```
initialize {SP_{n_i}(·), MTP(n_i), n_i.processed = false}_{i=1}^N
for k=1 to K do (S is an empty stack initially)
  for i=1 to N do
    n_i = "unprocessed" node with highest MTP
    push(S, n_i), n_i.processed = true
    perform single label pruning for node n_i
    for each "unprocessed" neighbor n_j do
      perform neighbor label pruning for node n_j
      transmit all messages m_ij(·) and update SP_{n_j}(·)
      update b_j(·), SP_{n_j}, MTP(n_j).
  for i = N downto 1 do
    node n_i = S.pop(), n_i.processed = false
    transmit all messages m_ij(·) to any "processed"
      neighbor n_j
    update MTP(n_j)
```

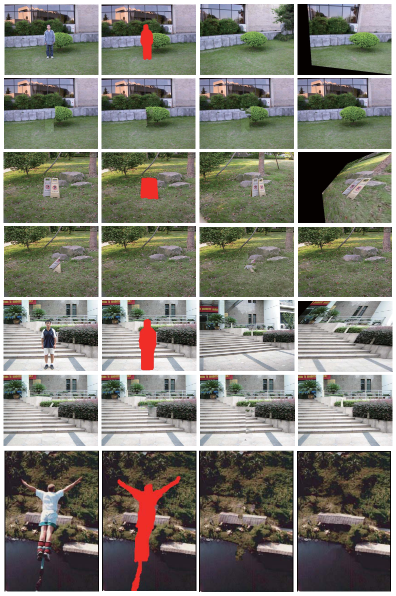

Figure 3: The first six rows are three examples of LDV completion, each example occupies two rows arranged like Fig. 1. The seventh row is an example of single image completion, from left to right: original image, masked image, result of Criminisi's method, and result of our method.

message from n_i to n_j will promote node n_j to choose $L_s(n_j)$ as its high belief labels in all likelihood as mentioned in Section 2.3. Suppose label $l_j \in L_s(n_j)$ then l_j would also have a good structural match with $l_{optimal}(n_i)$, thus its structural priority is at least $SP_{n_j}(l_j) = SP_{n_i} - S(l_{optimal}(n_i), l_j)$ according to (7). By this way, n_j inherits a high structural priority from n_i with a small difference $S(l_{optimal}(n_i), l_j)$. The high structural priority of node n_j leads to its high MTP, and similar situation will happen repeatedly to effectively encourage structure propagation.

3. EXPERIMENTAL RESULTS

We test SPBP on four LDV completion cases and one single image completion case. Note that SPBP can also be used for single image completion without neighbor label pruning. In Fig. 1, the man is removed manually and the unknown region (with 10134 pixels, marked red) occluded by him is restored. Apparently, restoring the occluded stone with complex shape is challenging. Single image completion method [3] can only "guess" the shape of the stone due to lack of necessary information and it gives a poor result also due to its greedy scheme. We perform the perspective transformation on LDV image using the homography matrix [11] to get a warped LDV image. A salient mismatch exists in the completion result of directly employing the warped LDV image since it reflects a scene with different depth. Previous LDV method [10] cannot work properly in this case since it is difficult to establish feature correspondence on the grassplot, and [11] also generates a poor result due to its greedy scheme. Thanks to dynamic weight of structural consistency and structural priority inheritance, the complex structure of the stone is faithfully restored by structure propagation with a correct priority, while texture regions such as grassplot are also completed well. In a similar way, Fig. 3 shows three examples of LDV completion with 8093, 8007 and 14218 unknown pixels respectively, and an example of single image completion with 7769 unknown pixels where the roof of the house (with salient structure) is restored much better than the greedy method in [3].

In our implementation, the patch size ranges from 5×5 to 11×11, the size of test images is about 472×354, the search window for defining labels of a node in LDV completion is ten times the size of the patch and the threshold D in neighbor label pruning is half of the size of search window. Due to the efficient neighbor label pruning, SPBP takes less than one minute on a 2.2GHz CPU to repair the large unknown regions shown in Fig. 1 and Fig. 3 for LDV completion.

4. CONCLUSION

This paper presents a new method called SPBP to deal with LDV completion. Two important extensions are made to Priority-BP: dynamic weight of structural consistency and structural priority inheritance so as to propagate structure with a correct priority while maintaining texture propagation in a global optimization scheme. Neighbor label pruning is developed to accelerate SPBP for LDV completion. Experimental results show that SPBP outperforms previous algorithms of LDV completion and it works well for single image completion.

Acknowledgement

This work is partially supported by National Basic Research Program of China (2009CB320802). We thank Ruanjie Wei in Shanghai Jiao Tong University for his effort in taking all photographs in our experiments and preparing illustration design.

5. REFERENCES

[1] Connelly Barnes, Eli Shechtman, Adam Finkelstein, Dan B Goldman. PatchMatch: A Randomized Correspondence Algorithm for Structural Image Editing. *ACM Trans. Graph.* 28(3), 2009.

[2] M. Bertalmio, G. Sapiro, V. Caselles, C. Ballester: Image inpainting. In *ACM SIGGRAPH*, pp. 417-424, 2000.

[3] A. Criminisi, P. Perez, K. Toyama. Region flling and object removal by exemplar-based image inpainting. *IEEE Trans. Image Processing*, 13(9), 2004.

[4] I. Drori,D. Cohen-or, H. Yeshurum: Fragment-based image completion. *ACM Trans. Graph.* 22(3):303-312, 2003.

[5] P. Felzenszwalb and D. Huttenlocher. Efficient Belief Propagation for Early Vision. *IJCV*, 70(1):41-54, 2006.

[6] James Hays, Alexei A. Efros. Scene completion using millions of photographs. *ACM Trans. Graph.* 26(3), 2007.

[7] T. Huang, et al. Image Inpainting by Global Structure and Texture Propagation. In *ACM Multimedia*, pp. 517-520, 2007.

[8] J.-Y. Jia, C.-K. Tang: Image repairing: Robust image synthesis by adaptive ND tensor voting. In *CVPR* pp. 643-650, 2003.

[9] N. Komodakis, G. Tziritas. Image Completion Using Global Optimization. In *CVPR*, 2006.

[10] C. Liu, Y. Guo, L. Pan, Q. Peng, F. Zhang : Image completion based on views of large displacement. *The Visual Computer* 23, 9-11: 833-841, 2007.

[11] C. Liu, Y. Yang, Q. Peng, J. Wang, W. Chen. Distortion Optimization based Image Completion from a Large Displacement View. *Computer Graphics Forum*, 27(7):1755-1764, 2008.

[12] C. Strecha, T.Tuytelaars, L. Van Gool. Dense matching of multiple wide-baseline views. In *ICCV* pp. 1194-1201, 2003.

[13] J. Sun, L. Yuan, J. Jia, H. Shum. Image completion with structure propagation. In *ACM SIGGRAPH*, pp. 861-868, 2005.

[14] Q. Wu, Y. Yu. Feature matching and deformation for texture synthesis. *ACM Trans. Graph.* 23(3):364-367, 2004.