

Transforming Image Completion

Alex Mansfield¹

<http://www.vision.ee.ethz.ch/~mansfiea/>

Mukta Prasad¹

<http://www.vision.ee.ethz.ch/~mprasad/>

Carsten Rother²

<http://research.microsoft.com/people/carrot/>

Toby Sharp²

<http://research.microsoft.com/people/tsharp/>

Pushmeet Kohli²

<http://research.microsoft.com/people/pkohli/>

Luc Van Gool¹

<http://www.vision.ee.ethz.ch/~vangool/>

¹ Computer Vision Laboratory

ETH Zürich

Switzerland

² Microsoft Research Ltd.

Cambridge

U.K.

Image completion is an important photo-editing task which involves synthetically filling a user-marked region (also called a hole) in an image, such that the image still appears natural to the user.

State-of-the-art image completion methods search for patches in the image that overlap consistently with existing nearby regions, locally match other image patches and produce an overall appearance that does not deviate much from regions elsewhere in the same image *e.g.* in [2, 4] or different images *e.g.* [3]. However, in many cases, the unknown region may never have been observed. For example, the part of the clock in the hole of Fig. 1 is not repeated in the image. However, it can be filled by rotating other parts of the clock. Similarly, noise, shading, perspective distortion and texture irregularities may force us to expand our training set by considering source patches in the larger space of natural transformations (scale, rotation, brightness change *etc.*) of the training set.

Our contribution is to extend image completion using the state-of-the-art method of Wexler *et al.* [4] to include transformations, and to investigate the best way to search for the best image patches and the optimal transformations to produce natural completions. We evaluate our work quantitatively as well as qualitatively in terms of the quality of results obtained.

Given an image S with a hole, we want to produce a completed target image T . A binary visibility map V takes a value 1 when a pixel is filled and 0 otherwise, as illustrated in Fig. 2. Our model is based on image patches, with the relative position of each pixel in the patch given by $\mathbf{i}_p \forall p$. We express the patch in S around \mathbf{n} compactly in vector form as $\mathbf{s}_n = \{S(\mathbf{n} + \mathbf{i}_p) | \forall p\}$, and similarly for patches \mathbf{t}_n and \mathbf{v}_n in T and V respectively.

For a transform parametrised by a set θ , the source patch is denoted as \mathbf{s}_{θ_n} . We extend [4] to formulate our problem as:

$$T^* = \operatorname{argmin}_T \left(\sum_n \min_{\theta_n} \|\mathbf{t}_n - \mathbf{s}_{\theta_n}\|^2 \right) \text{ subject to } V(\mathbf{i}) = 1 \quad \forall \mathbf{i}. \quad (1)$$

We consider *rotation* and *scale* as well as the usual *translation* as geometric transformations, and a *brightness shift* as a photometric transformation of a patch. This gives a total of five parameters in the set θ .

We optimise this model using the algorithm of [4], which iterates between two steps until convergence of the target image:

1. **Search:** For each patch in the current target image, the nearest neighbour patch in the source image is found. The result is a set of parameters θ_n for target patches around each pixel \mathbf{n} .
2. **Filling:** A new target image is generated based on the nearest neighbour patches.



Figure 1: Image completion with transformations can lead to much improved results over previous state of the art. The red area shows the hole.

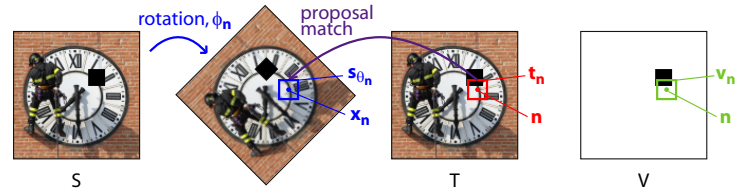


Figure 2: Our notation. For each patch in the target image T , we find matching patches in the source image S under transformations, *e.g.* rotation. The binary visibility map V specifies pixels still to be completed.

We consider and test different search methods. We consider discrete search over discrete variables and as an approximation for continuous variables, including the Patch Match algorithm [1]. We consider continuous optimisation with Levenberg-Marquadt for continuous variables. We also use a closed form search for the brightness shift.

The hole is filled by combining votes from the nearest neighbour source patches. The pixel colour taken is the median of all votes, found independently for each colour channel.

We perform many experiments to evaluate our different search methods and transformations. The true objective of image completion is to find a completion satisfactory to the user, so visual inspection is the natural test. However, we also evaluate our results quantitatively using the final energy and the RMS pixel colour error with respect to the ground truth, where possible. This error is a suitable measure of plausibility when the completion problem is sufficiently well constrained. Note that the energy function depends on many parameters; however, comparisons can always be made using the error, where the ground truth is available.

We show that including transformations improves a model for completion of images which exhibit redundancy over such transformations. Optimising this model is difficult, but the use of continuous search over the transformation parameters brings significant improvement in accuracy compared to previous search methods, with a good approximation achieved in much less time by the Patch Match based search algorithms.

We show that it is advantageous to limit the transformations considered to those that are necessary, to limit the addition of local minima. Finally, we show that the patch size is an important parameter that must be chosen in accordance with the problem image and the transformations considered.

All results were created with our own implementations, which are available under GNU General Public License at <http://www.vision.ee.ethz.ch/~mansfiea/ttransformic/>

- [1] C. Barnes, E. Shechtman, D. B. Goldman, and A. Finkelstein. The generalized patchmatch correspondence algorithm. In *Proc. ECCV*, 2010.
- [2] Alexei A. Efros and William T. Freeman. Image quilting for texture synthesis and transfer. *Proceedings of SIGGRAPH 2001*, pages 341–346, August 2001.
- [3] James Hays and Alexei A Efros. Scene completion using millions of photographs. *Proc. ACM SIGGRAPH*, 26(3), 2007.
- [4] Y. Wexler, E. Shechtman, and M. Irani. Space-time completion of video. *IEEE PAMI*, 29, 2007.