Interactive image completion with perspective correction

Abstract We present an interactive system for fragment-based image completion which exploits information about the approximate 3D structure in a scene in order to estimate and apply perspective corrections when copying a source fragment to a target position. Even though implicit 3D information is used, the interaction is strictly 2D, which makes the user interface very simple and intuitive. We propose different interaction metaphors in our system for providing 3D information interactively. Our search and matching procedure is done in the Fourier domain, and hence it is very fast and it allows us to use large fragments and multiple source images with high resolution while still obtaining interactive response times. Our image completion technique also takes user-specified structure information into account where we generalize the concept of feature curves to arbitrary sets of feature pixels. We demonstrate our technique on a number of difficult completion tasks.

Keywords Image completion · Image repair · Example-based synthesis · User interface

1 Introduction

Fragment-based image completion techniques are a very powerful tool to fill in missing pixel information, e.g., when removing a foreground object from a digital photo. The conceptual idea is to fill a hole in the image by copying small source fragments from known regions of the image such that they eventually completely cover the undefined region. The mutual overlap of the target fragments and the overlap between target fragments and the boundary of the hole is used to compute a similarity measure which controls the selection of the best source fragment candidate in order to guarantee a seamless appearance of the completed image.

The various approaches to fragment-based image completion mainly differ in three aspects. First, in the definition of the search space. Source fragments can be taken anywhere from the source image or only from certain user-defined sub-regions in order to preserve specific feature information or to reduce time complexity. Second, the selection of the best source fragment is based on a similarity measure, which can use pixel color information as well as structural information such as the presence and orientation of image features. Third, once the best source fragment is found, it has to be transformed to the target location in the image. Besides mere translation, certain types of affine transforms, like scaling and rotation, have been proposed.

The fundamental assumption, which justifies the fragment-based image completion approach, is that for a small enough image fragment, we can assume the scene, which is visible in this fragment, to be planar. Hence we can ignore all kinds of occlusion and dis-occlusion effects when copying a fragment from one image location to another, and therefore we do not need any true 3D information about the scene. However, the restriction to affine transforms of the fragments, as it has been done in the previous work, mathematically corresponds to the even more strict and somewhat unrealistic assumption that these planar scene fragments are aligned to the image plane. A natu-
Fig. 1a–e. The effect of perspective cor-
nexions for image completion. a shows
the input image. d shows the comple-
ction without using perspective correction
which leads to strong perspective artifacts.
When using the rectifying homography
defined in b and then performing the com-
pletion in the rectified space c we obtain
the solution in e.

eral generalization of the existing approaches is, therefore,
to allow for projective transforms of the image fragments,
which enables the compensation of perspective distortion
when the source and target scene fragments are not lying
in the same supporting 3D plane.

In this paper we present a new system for interac-
tive image completion that applies perspective corrections
when copying fragments. Through a simple interaction
metaphor the user can define a set of projective trans-
forms. Based on this information the system rectifies the
corresponding image regions and then performs the image
completion in rectified image space. An intuitive exten-
sion of this user interface allows the system to handle
even more complex scene geometries such as moderately
curved surfaces. An automatic snapping mechanism for
the points selected by the user guarantees continuity at the
boundaries of adjacent rectified regions. User-defined fea-
ture information is taken into account by encoding feature
proximity as an additional color channel to the image.

Figure 1 shows a simple example for an application of
our interaction metaphor and the resulting image comple-
ction which is free from perspective artifacts. In the results
Sect. 9 we present more examples, which demonstrate the
power and usefulness of our system.

2 Related work

There are two fundamental approaches to image comple-
tion: image inpainting methods and example-based ap-
proaches. Image inpainting [3, 5, 6, 28] is good in filling
small missing areas like thin gaps, e.g., when removing
scratches from old photographs, but the diffusion process
of image inpainting leads to blurring artifacts when trying
to complete large missing areas, e.g., after removing large
foreground objects from images.

Example-based approaches have their origin in tex-
ture synthesis, where large texture patches are synthe-
sized from small texture samples on a per-pixel basis using
a pixel neighborhood for color prediction [2, 14, 32]. In-
stead of this per-pixel procedure some approaches use
larger patches or fragments [13, 25] in order to speed up
the process. By using graph-cut techniques [24], optimal
boundaries between patches can be computed.

In the context of image completion there are some
approaches working on a per-pixel basis [7, 16, 20], but
fragment-based approaches [4, 9, 12, 31] usually produce
superior results. The automated method presented by
Drori et al. [12] leads to very impressive completions but
at the cost of high computation time. Our method achieves
results of comparable quality but it is up to two orders
of magnitude faster allowing for interactive control of the
process.

An important observation for image completion is that
the propagation of structure should be treated separately
from texture propagation [7, 9]. Recently Sun et al. [31]
presented an interactive approach, where feature curves
for structure propagation are supplied by the user. Putting
a human user into the loop usually leads to much more
plausible completions than the ones generated by sophis-
ticated heuristics. Hence user-control is considered an im-
portant feature. In our system the user also has the possi-
bility to specify structure information as additional color
channel to the image.

Other previous completion techniques take user-defined
points of interest [12] into account to control the propa-
gation direction when completing rotationally symmetric