Wavelet-Based Inpainting for Object Removal from Image Series

Sebastian Vetter, Marcin Grzegorzek, and Dietrich Paulus

Active Vision Group
Institute for Computational Visualistics
University of Koblenz-Landau
Universitätsstr. 1, 56070 Koblenz, Germany

Abstract. We propose several algorithmic extensions to inpainting that have been proposed to the spatial domain by other authors and apply them to an inpainting technique in the wavelet domain. We also introduce a new merging stage. We show how these techniques can be used to remove large objects in complex outdoor scenes automatically. We evaluate our approach quantitatively against the aforementioned inpainting methods and show that our extensions measurably increase the inpainting quality.

1 Introduction

The digital nature of most images that are used today, make it fairly easy to alter the image content using image processing tools. The removal of entire objects from images, however, is a very complicated task that usually requires the expertise of professionals to achieve sufficient results in the altered image. It is, therefore, an even greater challenge to automate this process such that the results are perceived as un-altered.

Our need for inpainting came from a project that required removal of objects, a colour checker in particular, from a series of highly structured images. Fig. 1 shows an example from the test set. In an initial processing step, the colour checker is located in the image and a binary mask is generated identifying the pixels belonging to this object. It is then removed using image inpainting.

We propose a modified approach for image inpainting based on an approach described by Ignácio et al. [1] using wavelet coefficients to determine the order in which the unknown region is filled. Their idea is adapted from Criminisi et al. [2] where the authors apply the same technique in the spatial domain. Our approach defines extensions to improve the wavelet-based inpainting approach by Ignácio et al. [1]. These extensions are similar to those proposed by Cheng et al. [3] to improve [2].

We outline related work in section 2, briefly introduce the idea of the wavelet-based inpainting approaches in section 3, describe our extensions to Ignácio et al. [1] in section 4 and provide the corresponding experimental results in section 5. We conclude in section 6.

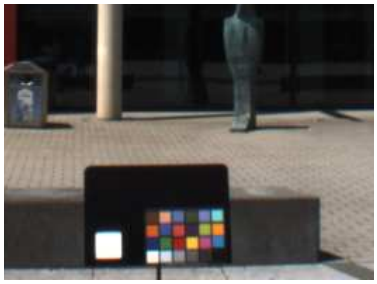



Fig. 1: Images from our test set showing the original (left) and the inpainted image using our approach with the colour checker removed (right)


Fig. 2: Some original scenes used in the test

2 Related Work

The removal of objects from still images has been the subject of many research projects. Initially, approaches only had to deal with missing data in the size of a few pixels. These damages could be repaired by using well-known interpolation techniques which are incorporated into almost any image processing application available.

Reconstructing larger regions of missing data, as they occur when removing objects from images is much more difficult. This task usually requires the expertise of professionals to provide a sufficient quality of results. Even more difficult is the development of automated algorithms that allow inpainting of such regions. Approaches in this field have slowly emerged from two fields of research, namely *texture synthesis* and *structure inpainting*. Research concerned with texture synthesis expects a small texture sample and tries to create a much larger texture image automatically. In contrast, structure inpainting initially focusses on the removal of small damages by reproducing intensities based on neighbouring pixels.

A popular attempt has been described by Harrison [4] who removes an object using exemplar-based texture synthesis, sampling from a given texture to fill the unknown intensities. This is closely related to the ideas of Garber and Sawchuk [5] as well as Efros and Leung [6] who proposed algorithms to create new inten-

sity values by sampling their best-matching neighbouring pixels. Wei and Levoy [7] extended the latter approach by combining a texture- and example-based technique working through different levels of coarseness in a pyramid scheme to circumvent the size restrictions of sample pixels that apply to [6].

Structure inpainting, in contrast, looks at the problem of filling in missing data from an artistic point of view. As part of their work, Bertalmio et al. [8] consolidated inpainting artists to examine their method of operation restoring a painting; they tried to replicate manual inpainting by propagating the known intensities into the unknown region along so called *isophotes*, representing the smallest spatial change of intensities, i.e. structure. The authors continued working on that problem and discovered that "different techniques work better for different parts" [9] and started to develop a new approach based on the decomposition of an image into two parts. One part represents the structure, the other part the texture of the damaged image and structure inpainting and texture synthesis are applied to the separated components. In a similar approach, Drori et al. [10] used adaptive circular fragments operating on different scales capturing both, global and local structures and approximating the missing region.

Criminisi et al. [2] proposed an approach employing rectangular patches that are iteratively filled depending on calculated priorities along the boundary of known and unknown region. Due to its intuitive principle, it was applied not only to object removal but also to remove rifts in puzzles of archaeological fragments as described in Sagiroglu and Ercil [11]. Cheng et al. [3] extended the priority equation of Criminisi et al. [2] and made it adjustable to the structural and textural characteristics specific to an individual image.

The simplicity of the concept in Criminisi et al [2] was taken up by Ignácio et al. [1] who applied it in the wavelet domain. They transform the image and binary mask and then use wavelet coefficients to determining the fill-order guided by a similarly defined priority.

3 Wavelet-Based Image Inpainting

Applying the approaches [2, 3, 1] to our images revealed that they all struggle with images of highly structured content such as buildings, pavements, etc. The extension described by Cheng et al. [3], however, shows an improvement compared to [2]. Further experiments revealed that results produced by [1] are of similar quality to those obtained by [3]. An example is shown in Fig. 3.

We observed that the results for [2] and [1] are largely dependent on the content of each individual image which is described Cheng et al. [3] as the ratio of structural and textural content. Their proposed extension to [2] incorporates these observations and provides an adjustable algorithm for inpainting images. In accordance with their observations, we decided to develop a similar extension to [1] to make it adjustable while exploiting its high-frequency coefficients. They provide the edge-related information suitable for correctly inpainting structures in highly complex images.

(a) Criminisi [2] (b) Cheng [3] (c) Ignácio [1]

Fig. 3: The images show the improvement achieved by Cheng et al. [3] towards the approach by Criminisi et al. [2] and Ignácio et al. [1].

We expect an input image I and a binary mask specifying the object to be removed resulting in a region of unknown intensities Ω. Samples to fill Ω are taken from Φ holding $I = \Omega \cup \Phi$. The filling algorithm operates on small image regions called *patches* where a patch is defined as a set of pixel locations $\Psi(\boldsymbol{p})$ centred at location $\boldsymbol{p}$. Each patch has a square size of $M \times M$ pixels.

In accordance to [1], we transform both, input image and binary mask, into the wavelet domain using a *Haar* wavelet and apply an iterative process of three consecutive steps in the wavelet domain: *determine the fill-order (A)*, *find the best match (B)* and *fill the search patch (C)*.

In step (A) we calculate a priority value for each patch $\Psi(\boldsymbol{p})$ that is centred on the boundary $\delta\Omega_W$ between the target region Ω_W and source region Φ_W in the transformed image. This is called the *fill-order*. In step (B) we use the search patch $\Psi(\hat{\boldsymbol{p}})$ with the highest priority and search for the most similar sample patch $\Psi(\hat{\boldsymbol{q}})$ in Φ_W providing the coefficients unknown in $\Psi(\hat{\boldsymbol{p}})$. In step (C) we fill $\Psi(\hat{\boldsymbol{p}})$ by simply copying the corresponding coefficients from $\Psi(\hat{\boldsymbol{q}})$.

After the entire target region Ω_W is filled with sampled coefficients we transform the result back into the spatial domain and obtain an image where Ω has been reconstructed by sampling from Φ.

4 Modified Wavelet Inpainting

In accordance to the original approach, the key concept of our approach remains its fill-order. The priority value determining the fill-order of Ω_W is calculated for each search patch $\Psi(\boldsymbol{p})$ centred on $\delta\Omega_W$. Ignácio et al. [1] define the priority value $P(\boldsymbol{p})$ as

$$P(\boldsymbol{p}) = K(\boldsymbol{p})\ S_S(\boldsymbol{p})\ S_O(\boldsymbol{p}) \quad , \tag{1}$$

where $K(\boldsymbol{p})$ denotes the confidence, $S_S(\boldsymbol{p})$ the structure significance, and $S_O(\boldsymbol{p})$ is the structure orientation at the pixel location $\boldsymbol{p}$.

The last two terms correspond to those named *edge strength* and *edge orientation* in [1]. We refrain from using these terms because the high frequencies in

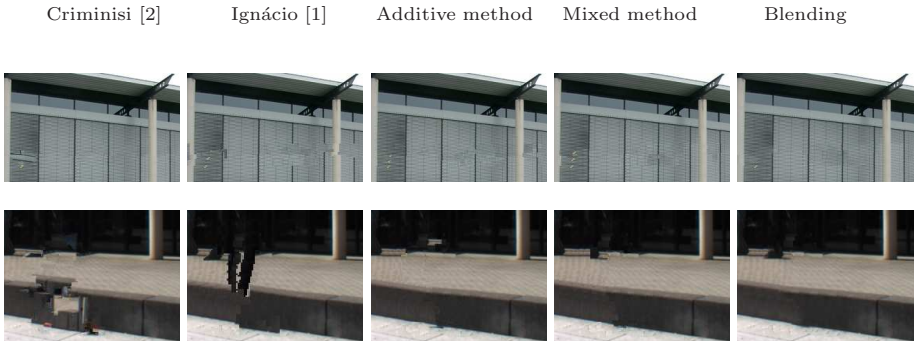


Fig. 4: Excerpts showing the inpainting results generated by [2] and [1] compared to our modifications. Each of our modifications show an increase in quality. Added blending produces the best results.

the wavelet domain are different from edge strength and edge orientation in the spatial domain as used in [2].

As in [1], we use the concept of a fill-order determined by priorities calculated for each search patch $\Psi(\boldsymbol{p})$ centred on $\delta\Omega_W$. We, however, introduce several modifications to the priority equations based on modifications to [2] proposed in [3]. We observed that corresponding modifications to equation 1 resulted in higher inpainting quality and, as a side effect, allowed adjusting the fill-order to the relative occurrence of structure and texture in an image. This allows us to counterbalance the high impact of $K(\boldsymbol{p})$ on $P(\boldsymbol{p})$ in [1] which resulted in a fill-order mainly driven by high reliability instead of structural properties, reducing the inpainting quality as shown in Fig. 4 where $K(\boldsymbol{p})$ outweighs the structural components and mainly inpaints the textures.

To evade this negative influence of $K(\boldsymbol{p})$, we propose to change the equation for $P(\boldsymbol{p})$ used in [1] to equation (2). This modification corresponds to the proposal of Cheng et al. in [3] to modify the priority equation in [2]. Fig. 5 illustrates the increase in variation of priorities. This makes the equation more resilent to noise and outliers in either component. We apply the same strategy to reduce the impact of the other components on $P(\boldsymbol{p})$. This leaves us with a modified equation for the priority obtained by adding the respective components. We define it as

$$P(\boldsymbol{p}) = K(\boldsymbol{p}) \; + \; S_S(\boldsymbol{p}) \; + \; S_O(\boldsymbol{p}) \tag{2}$$

4.1 Weighting the Components

The inpainting results respond immediately to these changes. Fig. 4 shows that inpainting the same image using the *Additive* approach increases the continuation of isophotes and reduces artefacts.

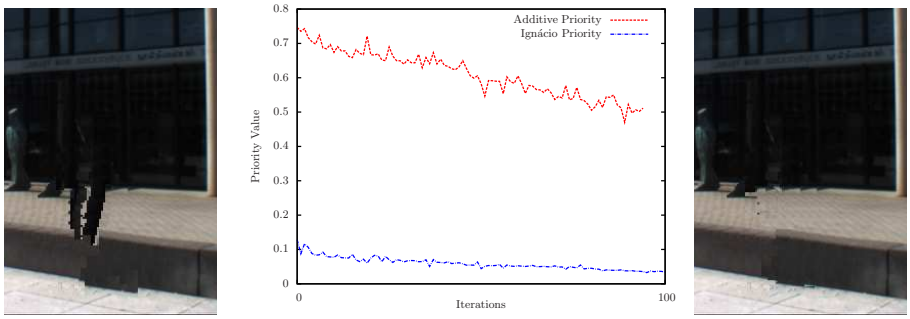


Fig. 5: The plot shows that priorities calculated using (2) vary more then priorities used in [1] which reduces the influence of noise and outliers. The difference in results is illustrated for Ignácio [1] (left) and our Additive approach (right).

This modified equation, as the original, applies the same structure-texture ratio to every image which is inaccurate for most images. This was already observed in [3] and inspired our next modification making equation (2) adjustable to the different characteristics of images.

We introduce the concept of adapting equation (2) by introducing two weighting parameters ω_S and ω_C to adjust the proportional dominance of the components. We name ω_S the *structure weight* and use it to control the relative importance of structure orientation and significance resulting in a weighted structure value $S(\boldsymbol{p})$ defined as

$$S(\boldsymbol{p}) = (1 - \omega_S)\ S_O(\boldsymbol{p})\ +\ \omega_S\ S_S(\boldsymbol{p}) \tag{3}$$

with $0 \leq \omega_S \leq 1$. Increasing ω_S stresses the structure significance S_S which resembles the strength of isophotes in a patch $\Psi(\boldsymbol{p})$. Reducing it boosts the orientation-related characteristic of isophotes disregarding their possible difference in significance.

We then weight the obtained $S(\boldsymbol{p})$ against the $K(\boldsymbol{p})$ using weight ω_C. This parameter controls the relative influence of confidence and structure components on the priority:

$$P_A(\boldsymbol{p}) = (1 - \omega_C)\ K(\boldsymbol{p})\ +\ \omega_C\ S(\boldsymbol{p}) \tag{4}$$

where $S(\boldsymbol{p})$ is the weighted structure value in equation 3 and ω_C holds $0 \leq \omega_C \leq 1$. In case the settings hold $\omega_S = \omega_C = 0.5$, the calculated priority corresponds to an unweighted Additive priority using a separate normalisation. Altering the fill-order by using a higher ω_C results in a dominance of the structural component where inpainting is mainly isophote-driven.

4.2 Reducing Adjustable Parameters

Experiments with this extension showed that finding suitable parameters ω_S and ω_C for equation (4) can be difficult and time consuming. This motivated

the reduction to a single weight ω_C combining the structure orientation and significance into a single component making ω_S obsolete. We call this new priority value *mixed priority* and define it as

$$P_M(\boldsymbol{p}) = (1 - \omega_C)K(\boldsymbol{p}) \; + \; \omega_C \left(S_O(\boldsymbol{p}) \cdot S_S(\boldsymbol{p})\right) \tag{5}$$

with $0 \leq \omega_C \leq 1$. Fig. 4 shows that results produced by *Mixed* priorities approach are similar to those produced using the Additive priorities. The most important isophotes are continued correctly and the remaining artefacts are very similar with the improvement of one parameter less.

4.3 Alpha Blending

So far we modified the fill-order by making it adjustable to image properties. Several visible artefacts remain: displaced isophotes, patch-shaped artefacts caused by adjoining textures and repetitive patterns in the background texture. To reduce these artefacts we intervene in the step of filling $\Psi(\hat{\boldsymbol{p}})$ at the end of each iteration and introduce a blending of coefficients into the filling process.

Manual inpainting tools in image processing applications, e.g. the *healing brush*, use a blending mask for the brush to provide a smooth transition between added intensities and the background. We adapted this concept and added a blending method to the filling of $\Psi(\hat{\boldsymbol{p}})$.

We use coefficients from $\Psi(\hat{\boldsymbol{q}})$ to fill $\Psi(\hat{\boldsymbol{p}})$ by copying corresponding values to Ω_W equivalent to [1]. Source coefficients in $\Psi(\hat{\boldsymbol{q}})$ however are now blended with corresponding coefficients from $\Psi(\hat{\boldsymbol{p}})$. We consider each pixel in the source region of $\Psi(\hat{\boldsymbol{p}})$ in the blending process weighting it against its corresponding coefficient in $\Psi(\hat{\boldsymbol{q}})$.

Let $\boldsymbol{r}_i$ be a pixel in $\Psi(\hat{\boldsymbol{p}})$ holding $\boldsymbol{r}_i \in \Phi_W \cap \Psi(\hat{\boldsymbol{p}})$ and its correspondence $\boldsymbol{s}_i \in \Psi(\hat{\boldsymbol{q}})$. The new coefficient $\boldsymbol{r}_i$ is calculated by weighting the current coefficient $\boldsymbol{r}_i$ with the one corresponding to $\boldsymbol{s}_i$. The weight is

$$\omega_G = \frac{1}{2\pi\sigma^2} \; \exp\left(-\frac{d_x^2 + d_y^2}{2\sigma^2}\right) \tag{6}$$

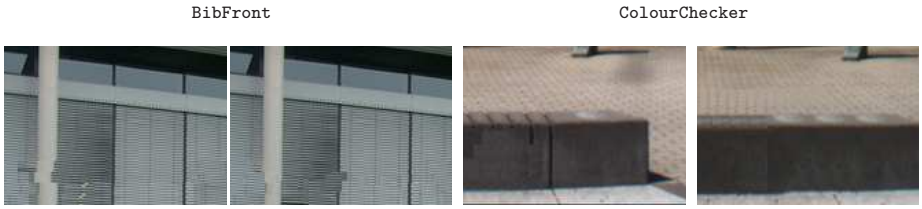
where d_x and d_y denote the distance of $\boldsymbol{s}_i$ to its patch centre in horizontal and vertical direction respectively. The weight ω_G is determined by a Gaussian distribution weighting each pair of coefficients. The blending is carried out accordingly for each subband which is defined as

$$W_n(\hat{\boldsymbol{r}}_i) = (1 - \omega_G) \cdot W_n(\boldsymbol{r}_i) \; + \; \omega_G \cdot W_n(\boldsymbol{s}_i) \tag{7}$$

where $n = a, v, h, d$ correspond to the four wavelet subbands. ω_G holds $0 \leq \omega_G \leq 1$ to ensure that it does not exceed 1, hence, retaining its original energy.

5 Experiments

We applied these modifications to our campus images (Fig. 2) and examined the inpainting results. As the example in Fig. 4 shows, our first modification



(a) Additive (b) Mixed (c) Additive (d) Mixed

Fig. 6: Samples inpainted with the Mixed approach, shown in (b) and (d), yield better results compared to the Additive modifications, shown in (a) and (c).

proposing two parameters to adjust the fill-order to the image-specific properties performs much better then the original approach which introduces visible artefacts. Our adjustable approach adapts to the characteristics of the image and increases perceived quality.

As an alternative, we proposed the *Mixed* approach in section 4.2 reducing the complexity of finding the correct set of parameters producing results equivalent to those of the Additive approach. Fig. 6 illustrates this with slight improvements using the Mixed approach.

We also introduced a blended filling process at the end of each iteration. The results in Fig. 4 show that using our blended approach reduces the amount of artefacts introduced by patches, improves the fill-order and smooths the transition of different textures.

Hence, the most reliable modification providing the best quality of results is the Mixed approach extended by our blending method. Its reduced parametric complexity with similar or even improved results, compared to the other approaches, makes it the favoured combination of our extensions. Table 1 shows that our extensions provide improved results compare to Ignácio et al. [1] and Criminisi et al. [2] and provides similar to better results when compared to Cheng et al. [3]. In addition, these result show that our extension provides a more consistent quality for all images whereas they can vary extensibly for the other approaches.

6 Conclusion

We conclude that our Mixed approach with blended filling increases quality for highly structured images. We achieved a higher stability with an improved fill-order and increase the inpainting quality adjusting the weight in equation (5) to the characteristics of each image individually. The optimal weights, however, have to be retrieved empirically.

We state that there is no generic solution to automatically remove an object from an image. Although the content of our images are very similar in terms of their texture-structure ratio, the remaining differences still require an adjustable

Table 1: PSNR for selected images from the series inpainted with [2], [3], [1] and our modifications.

	PSNR (dB)					
Image	Criminisi [2]	Cheng [3]	Ignácio [1]	Additive	Mixed	Blended
Entrance	44.62	49.17	46.45	46.73	46.67	47.37
G_Entrance_Top	44.61	48.72	47.45	47.34	46.79	48.11
Drain	44.26	44.52	43.89	44.15	43.97	44.52
LibFront_HorizLong	40.93	39.82	39.99	39.53	40.30	41.60
Shutter	31.23	40.55	40.25	42.28	42.56	42.74
Gravel	35.30	34.74	35.50	35.85	35.85	36.32
Menseria	42.55	43.76	41.91	42.39	42.56	43.04
Menseria_Top	36.48	28.90	33.60	35.61	35.97	36.80
Obelisk	29.66	29.38	33.74	34.01	34.03	34.17
CampusWater	31.889	37.05	35.06	34.61	34.88	36.31

solution to provide sufficient results. An approach proposing an automated solution, therefore, depends on the identification and classification of different image properties for an automated adjustment of the described parameters.

References

1. Ignacio, U.A., Jung, C.R.: Block-based image inpainting in the wavelet domain. The Visual Computer **23**(9-11) (2007)
2. Criminisi, A., Perez, P., Toyama, K.: Region filling and object removal by exemplar-based image inpainting. IEEE Transactions on Image Processing **13**(9) (2004)
3. Cheng, W.H., Hsieh, C.W., Lin, S.K., Wang, C.W., Wu, J.L.: Robust algorithm for exemplar-based image inpainting. In: Proceedings of CGIV. (2005) 64–69
4. Harrison, P.: A non-hierarchical procedure for re-synthesis of complex textures. In: WSCG. (2001) 190–197
5. Garber, D.D., Sawchuk, A.A.: Computational models for texture analysis and synthesis. In: Image Understanding Workshop. (1981) 69–88
6. Efros, A.A., Leung, T.K.: Texture synthesis by non-parametric sampling. In: ICCV. (1999) 1033–1038
7. Wei, L.Y., Levoy, M.: Fast texture synthesis using tree-structured vector quantization. In: SIGGraph-00. (2000) 479–488
8. Bertalmio, M., Sapiro, G., Caselles, V., Ballester, C.: Image inpainting. In: Proceedings of SIGGRAPH. (2000)
9. Bertalmio, M., Vese, L., Sapiro, G., Osher, S.: Simultaneous structure and texture image inpainting. IEEE Transactions on Image Processing **12** (2003) 882–889
10. Drori, I., Cohen-Or, D., Yeshurum, H.: Fragment-based image completion. In: Proceedings of ACM SIGGRAPH 2003. Volume 22. (2003) 303 – 312
11. Sagiroglu, M.S., Ercil, A.: A texture based matching approach for automated assembly of puzzles. In: ICPR. (2006) III: 1036–1041

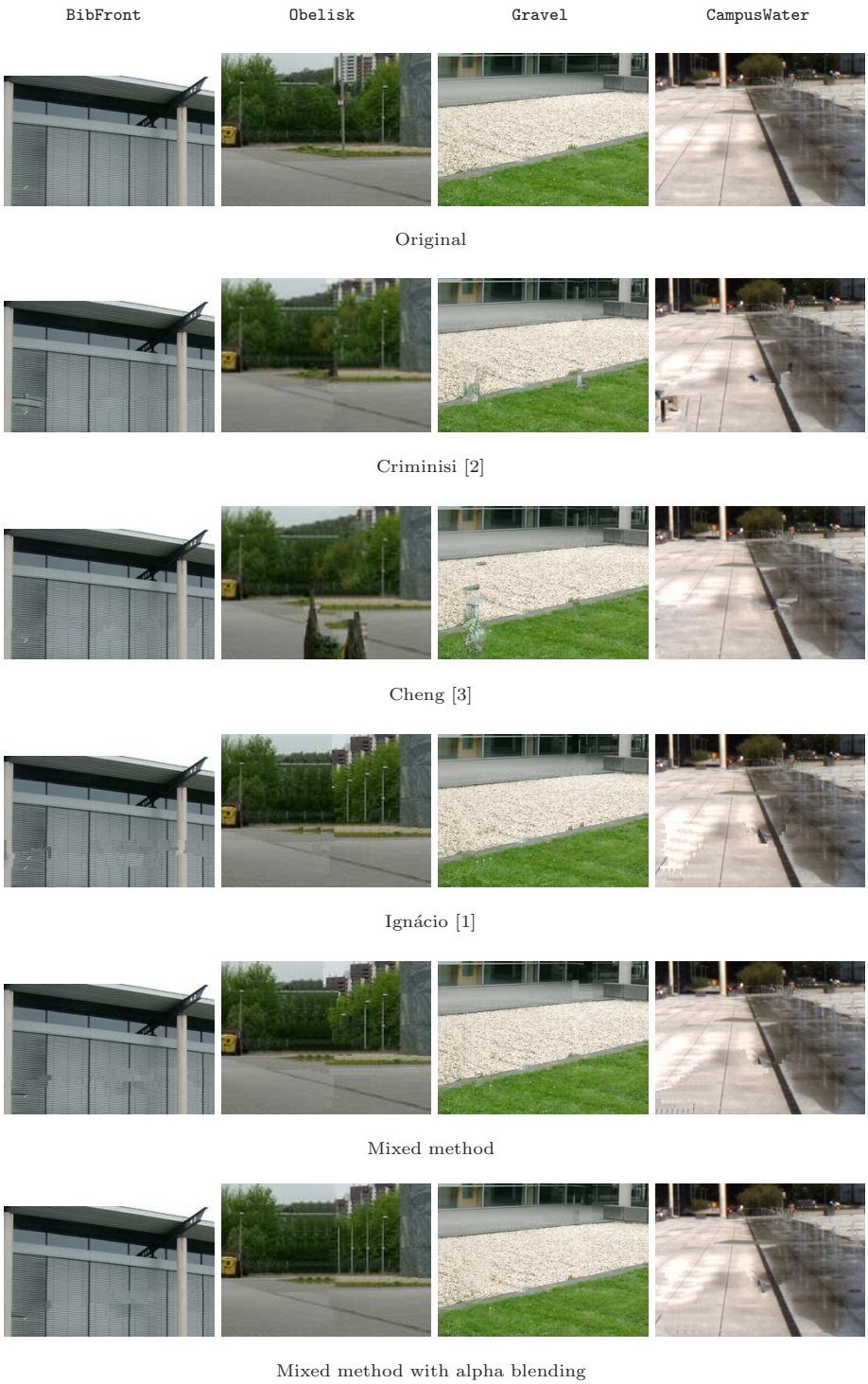


Fig. 7: Excerpts showing the inpainting results generated by [2] and [1] compared to our modifications. Each of our modifications show an increase in quality. Added blending produces the best results.