Transfer of Albedo and Local Depth Variation to Photo-Textures

Francho Melendez
Loughborough University
F.A.Melendez@lboro.ac.uk

Mashhuda Glencross
Loughborough University
M.Glencross@lboro.ac.uk

Jonathan Starck
The Foundry Visionmongers
jon.starck@thefoundry.co.uk

Gregory J. Ward
Dolby Laboratories
gward@dolby.com

ABSTRACT

Acquisition of displacement and albedo maps for full building façades is a difficult problem and traditionally achieved through a labor intensive artistic process.

In this paper, we present a material appearance transfer method, Transfer by Analogy, designed to infer surface detail and diffuse reflectance for textured surfaces like the present in building façades. We begin by acquiring small *exemplars* (displacement and albedo maps), in accessible areas, where capture conditions can be controlled. We then transfer these properties to a complete photo-texture constructed from reference images and captured under diffuse daylight illumination.

Our approach allows super-resolution inference of albedo and displacement from information in the photo-texture. When transferring appearance from multiple exemplars to façades containing multiple materials, our approach also sidesteps the need for segmentation.

We show how we use these methods to create relightable models with a high degree of texture detail, reproducing the visually rich self-shadowing effects that would normally be difficult to capture using just simple consumer equipment.

Categories and Subject Descriptors

I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and RealismColor, shading, shadowing, and texture

Keywords

Texture Transfer, Albedo, Displacement Map, 3D Reconstruction

1. INTRODUCTION

In this paper we present a semi-automatic method to produce displacement (*meso-structure*) and reflectance (*albedo*) maps, for visually rich textures, suitable for creating relightable 3D building models. Techniques to infer material appearance for use in making 3D assets are valuable for computer games, film post-production, archeology and architectural visualization. These applications need to be capable of rendering the recovered scenes under changing lighting as well as view point and require an increasing level of detail and realism. Material appearance is often conveyed by texture alone (recovered from images), but this appearance is only valid under the originally photographed viewing and lighting conditions. Realistic relighting requires a model that represents surface geometry and reflectance characteristics. A vast body of literature exists on methods for capturing reflectance and geometry, but capturing this information for large outdoor urban scenes, remains a difficult problem. Such scenes in particular offer interesting challenges since controlling lighting and access to areas to obtain suitable views rapidly becomes impractical. A commonly employed approach is for artists to manually create displacement and albedo maps, using a photo-texture as a reference.

To address this problem, we present a technique designed for transferring visually high-quality material appearance, captured at close range using standard digital SLR equipment, to a photo-texture. Material appearance, in this case, is constrained to diffuse albedo and shading information.

Employing a transfer approach rather than a synthesis approach preserves the original structure and appearance contained in the photo-texture. Our transfer approaches represent a good trade-off between simplicity of data capture and the quality of the results achieved for a range of broadly Lambertian materials used in construction.

The main advantages of the method presented here are:

- A simple capture process reduced to minimum equipment and calibration.
- The method is able to increase the resolution of the original photo-texture using information captured at close range in the exemplars.
- The algorithm automates the segmentation of materials present in the photo-texture and the associations with exemplars.
- In this paper we also provide a practical solution for addressing differences in lighting conditions between exemplars and the target photo-texture.

The remainder of this paper is organized as a brief outline of related work, a detailed description of our material appearance transfer approach, a an evaluation of its capabilities, results of the transfer applied to full building façades, future work and concluding remarks.

2. RELATED WORK

We begin by first giving a brief overview of these, structured in two sections: the state of the art in reflectance and shape estimation from images; and texture synthesis and transfer techniques.

CVMP '12, December 5 – 6, 2012, London, United Kingdom

2.1 Image Based Reflectance and Shape Recovery

The reflectance properties of opaque surfaces have been captured and modeled in a variety of different ways; we refer the reader to Dorsey et al. [10] for a thorough treatment of the topic. The Bidirectional Reflectance Distribution Function (BRDF) can be measured from images captured under a range of different viewing and lighting conditions [31, 24, 23, 29]. Capturing similar data for textured surfaces enables the creation of the Bidirectional Texture Function (BTF) [7]. Researchers have built upon the idea of the use of multiple light sources (photometric stereo) [36] to capture material appearance from fewer view points, effectively recovering albedo and local surface orientation [28, 12]. Paterson et al's [26] material capture approach combines photometric stereo with multiple view geometry captures to recover displacement maps and inhomogeneous BRDFs over nearly planar samples. Ward and Glencross [32] employ a similar approach to estimate albedo (diffuse reflectance), based on single view multi-flash captures in conjunction with shape from shading [16, 21, 39, 13].

For large surfaces, a possible solution often employed is to combine laser-scans of a scene, together with sky captures using an incident illumination measuring device and representative BRDF samples [8]. This data can be used within the inverse rendering framework [9, 4] to estimate the reflectance properties of a large complex scene. However, this approach requires specialized equipment and careful data collection. Without measuring the lighting for every image captured in the scene, further assumptions have to be made. Yu et al. estimated two pseudo-BRDFs for a polygonal model by fitting a small number of photographs, captured under clear sky conditions, to a parameterized model of the sky [38]. This approach offers improved relightable textures compared to using the original images, however no surface detail is captured and at least two lighting conditions per texture are required.

An interesting approach was also presented by Xu et al. [37] where the ratio between the green channel of the capture images and the reflected laser intensity is used to correct all color channels.

Rather than trying to explicitly solve the inverse rendering problem, which is ill-posed without capturing lighting and accurate geometry, we approach this problem from the texture transfer perspective. We propose capturing material appearance models of samples (exemplars) where we can control lighting conditions and then extrapolate these properties to the rest of the surface. We focus on two essential characteristics of the texture: albedo and meso-structure (depth). Glencross et al. [13] showed through evaluation with human subjects that this provides enough information to produce perceptually plausible relightable models of a great variety of textures.

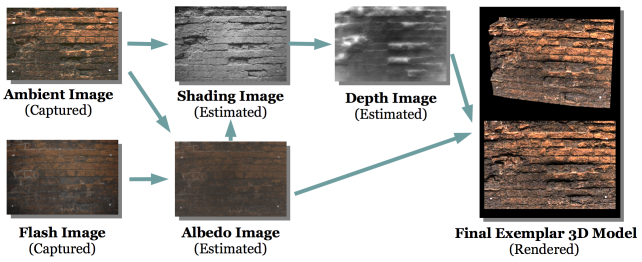


Figure 1: The capture process for Surface Depth Hallucination.

We capture our stimuli exemplars using Surface Depth Hallucination which is summarized in Figure 1. This technique requires capturing a photograph of opaque surfaces under natural diffuse lighting (ambient image), ideally on a cloudy day thus avoiding hard shadows, and another one firing a flash (flash image). By subtracting the ambient image from the flash image, and dividing by a calibration image, we compute an approximate albedo map. A shading image is calculated as the ratio of the ambient image and the albedo image, and used to create a depth map through a per-pixel dark-is-deep approach. A relightable 3D model can be created and rendered from the depth map and the albedo map. Due to flash guide distance limitations, exemplars captured using this method must be restricted to small surfaces (around one square meter), in order to keep detail in cracks and crevices sufficiently illuminated.

2.2 Texture Synthesis and Transfer

Over the last decade texture synthesis by exemplar has been a very active area of research. We refer the reader to state of the art reports for a complete review [20, 33]. This idea has been demonstrated to work effectively for synthesizing a wide variety of textures. New pixels or patches are generated by choosing the best candidate from a given exemplar, such that it is coherent with the already synthesized texture. For globally-varying textures, a control map is often used to drive this type of synthesis. Ashikhmin [1] synthesized an output conditioned by a colored map, and similar ideas are used in patch-based synthesis, defining the concept of texture transfer [11]. An extension of this idea, is the notion of Image Analogies [15].

Histogram matching was explored for synthesizing stochastic textures [14] and detail in textures [17] by matching the histogram of noise patterns to a texture sample. Glencross et al. [13] extended this work to transfer albedo and shading to similar sized surface samples. Melendez et al. [25] described a pipeline to apply this process to full building façades. Their method requires exemplar and target image to be statistically similar. For example, if we have a brick wall with moss, the amount of moss should be similar in proportion to the amount of brick in both exemplar and target image. Our novel approach can be included in the pipeline of Melendez el at., but since it matches materials localy, it overcomes the necesity of a good match in global statistics. This also allows us to run the transfer against a set of exemplars, without the necesity of segmenting the photo-texture and manually providing associations. Our method also allows for super-resolution to ensure scalability of our results to larger surfaces. Normally a photo-texture for a building is captured from far away, and therefore inference of high quality surface detail is severely limited by the resolution of the texture map.

We draw inspiration from the Image Analogies algorithm to define both albedo estimation and depth inference as a multi-exemplar based texture transfer problem.

3. MATERIAL TRANSFER

We consider the problem of acquiring albedo and depth maps for a photo-texture as a material transfer problem from image-based exemplars. In this context, we identify a material with a texture that has specific characteristics in terms of its reflectance and geometric structure, for example a type of brick wall. This would include the brick, the mortar, and even the dirt and other texture variations.

Using this definition of a material, Surface Depth Hallucination allows us to capture valid image-based exemplars in the form of albedo and depth maps, for texture reflectance and meso-structure. Our proposed transfer techniques facilitate the creation of the corresponding albedo and depth maps for large surfaces, such as a building façade, from a photo-texture and the previously captured exemplars. We capture a texture map under natural diffuse lighting conditions (no flash needed) for the complete façade and use this as a guide to transfer albedo and depth from representative exemplars.

3.1 Problem Statement

To illustrate the problem and evaluate our transfer technique, we use pairs of exemplars of the same texture as stimuli in which the scale and lighting conditions are consistent. This allows us to quantify the quality of the transfer process by comparing our results with the captured maps considered as ground-truth for perceptually plausible relightable models.

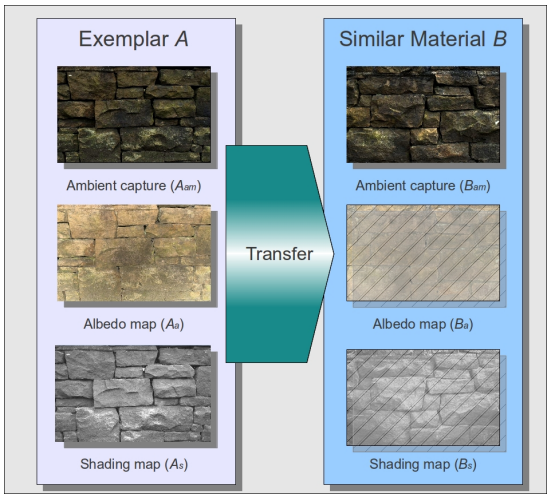


Figure 2: Material Transfer Problem: Using an exemplar A, generate new albedo and shading maps for a new ambient image B

The material transfer process is described in Figure 2. A material M is defined by an exemplar A which is in turn composed from three maps: an ambient map (A_{am}), a shading map (A_s), and an albedo map (A_a). Now consider another sample B of the same material M, for which only the ambient map is available. The aim is to synthesize a new shading map (B_s) and albedo map (B_a), from B_{am} and the exemplar A.

The ambient capture contains shadowing due to the geometry of the texture, and also color shifts due to natural lighting. From this information alone, resolving the ambiguity between albedo and shading is an ill-posed problem. We use an exemplar with similar characteristics where this ambiguity has been solved to help us arrive at a good approximation of both albedo and shading of the new image.

3.2 Transfer by Analogy

We propose using a locally adaptable transfer approach inspired by the idea of Image Analogies [15]. This method takes as input an image and a filtered version of it, and a target image to which we want to apply the same filter. Applying this terminology to our case, given an unfiltered source image A_{am} and two filtered source images A_a, and A_s, along with an additional unfiltered target image B_{am}, the aim is to synthesize two new filtered target images B_a and B_s.

This idea is illustrated in Figure 3. By comparing A_{am} and B_{am}, we find the patch in the exemplar A_{am} that best matches the appearance of every patch in the target image (B_{am}). Taking the same patch from the albedo and shading maps (A_a,A_s), we can create a new albedo and shading map for the new image (B_a,B_s). We define a patch for every pixel which can be understood as a descriptor for this pixel. The *best match* for a pixel will be the one with the most similar descriptors for a given metric.

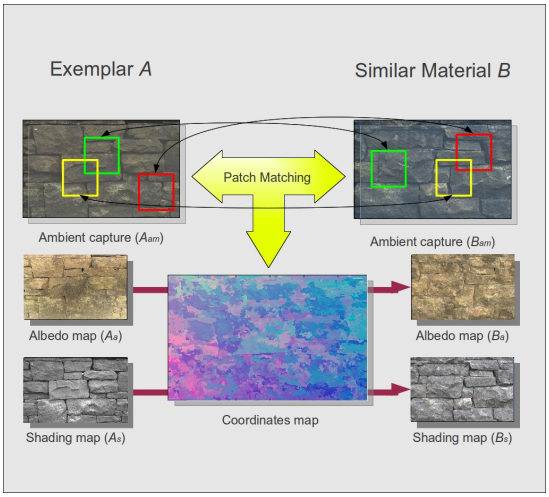


Figure 3: Trasnfer by Analogy.

3.2.1 Pixel Descriptor

The use of patches as descriptors is a common approach for exemplar-based texture synthesis [33]. In our experiments, we achieved the best results with $7x7$ patches using the RGB channels. This also provides a good trade-off between complexity of the descriptor and execution time. However, the optimal patch size can vary depending on the feature size of the texture, as it happens with most patch-based texture synthesis algorithms.

We add an extra channel to the RGB containing a distance-to-feature mask similar to the one used by Lefebvre and Hoppe [22]. Feature masks, and distance to feature masks, are used in texture synthesis to help the new synthesized texture to preserve the spatial structure of the exemplar. This is generated by computing the *Signed Distance Field* of a binary feature mask which is computed automatically from the ambient images A_{am} and B_{am}. The binary mask is the result of dividing the grey scale version of the image by a blurred version of itself, efectively extracting the high frequencies, and then thresholding it to create a binary image.

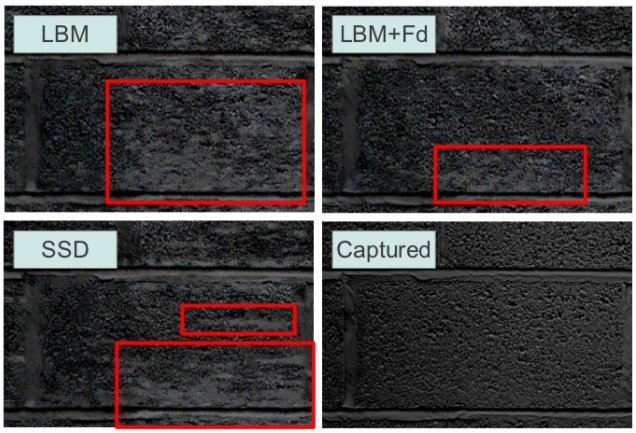


Figure 4: Close-up detail for shading images. LBM and SSD have gray areas where high-frequency detail is lost, while LBM+Fd better preserves the detail.

We observed that the use of a mask with this characteristic improves the transfer of high frequency detail, especially in the case

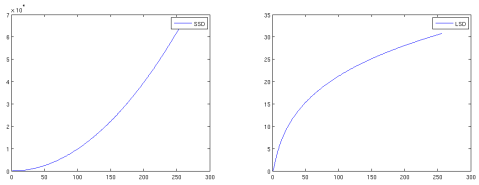
of the shading map. In Figure 4 we show this effect on a texture with high frequency detail using different metrics to match descriptors. Both sets of results based only on RGB channels, *Sum of square differences* (SSD) and our novel Log-Based Metric (LBM) (described in Section 3.2.2), contain areas with almost no high frequency detail (inside the red rectangles). Adding our distance to feature mask (LBM+Fd) clearly recovers more detail in the resulting shading image in the areas where the other metrics fail.

3.2.2 Appearance Metric

A second factor for the transfer is the metric or distance between descriptors used. SSD is the most common metric used in patch-based texture synthesis. Empirically we found that a novel LBM presented in Equation 1 can provide a better structural coherence than SSD for some textures.

$$LBM = \sum_{x,y} \log(1 + abs(x - y))^2 \quad (1)$$

The idea behind this metric is to limit the penalty for pixels that are very different. Figure 5 shows the profiles of SSD and our LBM for one pixel. The penalty increases rapidly but beyond a certain point it tends towards levelling-off for large pixel differences.



(a) SSD (b) LBM

Figure 5: Profile of penalty for difference on the compared pixel.

It is important to note that there are $7 \times 7 \times 3$ values to be compared and added. This metric also makes the penalty smaller per pixel. Consequently, more *bad* pixel matches are needed to penalize a good match. For example, a patch containing 49 pixels and one channel, where all the pixels match perfectly except for one which has the maximum difference $x = 255$, would have an SSD equal to $x^2 = 65025$ and a LBM of $\log(x+1)^2 = 5.799$. On the other hand, a patch with all it's pixels with difference $x_p = 36$, would have an SSD equal to $\sum_p x_p^2 = 65024$ and a LBM of $\sum_p \log(1+x_p)^2 = 120.50$. Our new metric therefore selects good global matches and dismisses large local errors in opposition to SSD. The local errors are compensated for when reconstructing the final image from the patches, (see Equation 2) where pixels are averaged according to the local error.

We evaluate our metrics against the sum of squared differences by computing the mean error between the transferred maps and the exemplars. Also we compare the results with and without using the feature mask, and the effect of giving more importance to the feature mask (LBM+Fd2). Figure 6 shows the computed average percentage error against the maximum possible error for a number of tested metrics. Our LBM shows consistently lower error. Although the numerical improvement is limited, we observed that when applied to several materials, it maintains better spatial consistency, resulting in better association of the correct material. We discuss this process in more detail in Section 5.

With this metric and descriptor, we have defined the *best match* between pixels of two images. In the next section, we present how

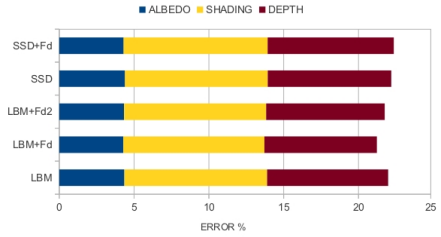


Figure 6: Comparison between metrics. SSD+Fd: Sum of Square Differences with Feature Distance Mask; SSD: Sum of Square Differences without Mask; LBM+Fd2: Log-Based Metric with Feature Distance Mask Squared; LMB+Fd: Log-Based Metric with Feature Distance Mask; LBM: Log-Based Metric without Mask.

to efficiently find the best match and how to create the final albedo and shading maps for the target image.

3.2.3 Approximate Nearest Neighbor Search

We find the best match by performing a nearest neighbor search. The problem of finding the nearest neighboring patch in a medium / large sized image rapidly becomes computationally expensive. Since patch-based sampling methods have become popular for image and video synthesis, many researchers have studied optimizing this process [34, 35, 18, 19]. In our implementation, we employ a recent patch matching algorithm from Barnes et al. [2] that performs at interactive rates for their application. This algorithm begins with a random initialization, and then uses an iterative process consisting of two steps: a propagation that searches within the previously matched pixels, and a random search to avoid local minima. This process proceeds in scan order (from left to right, top to bottom) for odd iterations, and in the inverse order for even iterations. Since the algorithm converges quickly to a solution, it typically provides a good level of detail within 5-10 iterations. We perform our matching in a coarse-to-fine grain fashion, as described in Figure 7, to improve the chances of converging to a globally optimal solution using 10 iterations per level.

```
PROCEDURE TransferByAnalogy(A_am, A_s, A_a, B_am, &B_s, &B_a)
{
   Compute_Gaussian_Pyramids(A_am);
   Compute_Gaussian_Pyramids(B_am);
   FOR EACH (level L)// from coarsest to finest
   {
      IF ( L == coarsestLevel )
         InitialisazeRandom(coordinatesMap);
      FOR iteration IN Number_of_Iterations
         {
            FOR EACH (pixel(q, iteration)) IN B_am //scaline order or inverse order
            {                                      // depending on the iteration
               p_old ¬ coordinatesMap(q);
               p_new1 ¬ Propagation(q, coordinatesMap, A_am, B_am);
               p_new2 ¬ RandomSearch(q, coordinatesMap, A_am, B_am);
               p ¬ best_of(p_old, p_new1, pnew_2);
               coordinatesMap(q) ¬ p;
            }
      }
      scaleUP(coordinatesMap);
   }
   reconstructFromCoordinates(coordinatesMap, A_a, &B_a) ;
   reconstructFromCoordinates(coordinatesMap, A_s, &B_s) ;
}
```

Figure 7: Fast Transfer by Analogy Algorithm.

In comparison with the original *Image Analogies* algorithm, we have removed the synthesis step, since we noticed in our experiments that this step can remove original features from the target texture, expecially when such features are not present in the exem-

plar, and our objective is to preserve the original texture.
The output of this stage of our algorithm is a coordinate map that associates every pixel in the target image with a pixel in the exemplar. We use this information to reconstruct the albedo and shading maps from the exemplar's corresponding maps. Simply taking the pixel indicated by the coordinate map introduces unnatural looking artifacts as can be seen in Figure 8(left). In contrast, we use a weighted average using pixel voting [30]. Pixel voting produces smoother transitions than simply using the best matched pixel (direct reconstruction) which tend to create pixelated images, and corrects for potential local mis-matching within patches (visible when zoomed in). It also creates sharper results than averaging (see comparison in Figure 8 (middle and right)). This is important to keep the structure of the texture coherent, especially in the shading image where noise can produce undesirable high frequencies in the resulting geometry.

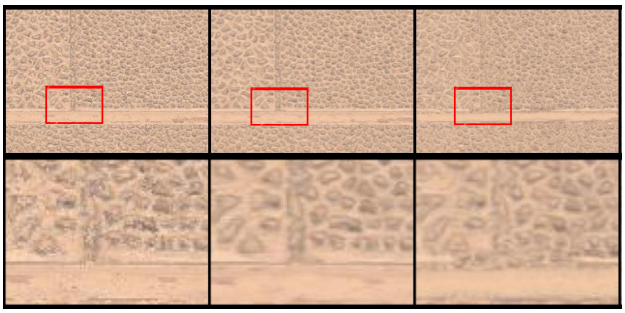
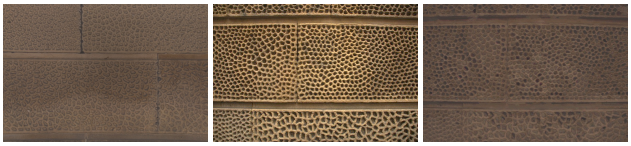


Figure 8: Recovered albedo image. Left: reconstructed with direct reconstruction from the coordinate map. Middle: using pixel voting. Right: using equal average.

Since our matching is performed in patches, every pixel is contained in as many patches as there are pixels within a patch. The weighting is described by Equation 2.

$$p(x,y)_{f_{new}} = \sum_n (p(x,y)_f^n * \frac{(1 - SD(p(x,y)_{am}^n - p(x,y)_{am}))}{\sum_n (1 - SD(p(x,y)_{am}^n - p(x,y)_{am})}) \quad (2)$$

For every pixel $p(x,y)$ we compute the value of the filtered image f_{new} as a weighted average of all the pixels p_f^n falling within the coordinates x,y from the different patches of the filtered image. For a 7×7 patch, every pixel is present in $n = 49$ patches. The weighting metric is computed as inversely proportional to the error (squared difference (SD)) which is normalized to sum of square differences of all the pixels in $p(x,y)$. Note that these differences are computed using the unfiltered image (the same used for finding the best corresponding patch) which we denote by the subscript $_{am}$ for *ambient*. This way pixels that match better contribute more to the result. This local combination of the values helps to eliminate pixels that match poorly, preserving the good matches. This combines well with our LBM, which prefers patch matching allowing for big differences but in a small proportion of the pixels in a patch.
Figure 9 shows the captured albedo map of an exemplar A_a, then a photograph under diffuse lighting of another sample of the same texture B_{am}. Finally the transferred albedo map for $B_{a'}$, wich is being reconstructed from patches of A_a. Figure 10 shows the resulting model, using the transferred albedo and shading maps, compared with captured exemplar used as the ground-truth. The transferred model compares well with the original and the produced albedo and geometry provide a good level of detail.



(a) A_a (b) B_{am} (c) $B_{a'}$(Transferred)

Figure 9: Transfer by Analogy creates a new albedo map with the same appearance as the captured A, but keeping the structure of the new sample B.

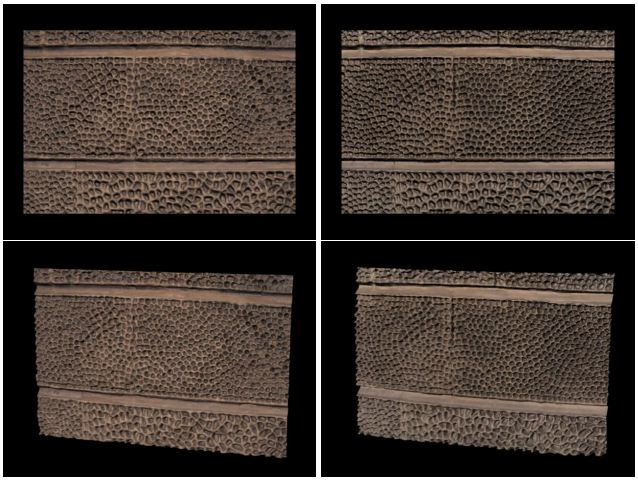


Figure 10: Transferred Model using Transfer by Analogy (left) and captured using Surface Depth Hallucination (right).

4. EVALUATION

We perform a series of experiments to evaluate the transfer capabilities of our method. We capture pairs of exemplars of the same texture, for example, A and B are two different exemplars of the same red brick wall. We transfer the characteristics from the exemplar A to the ambient capture of B (B_{am}) creating albedo B'_a and shading B'_s maps. These are compared to the captured maps B_a and B_s. We can also compare the result of transferring the characteristics of B to A, so we get two results per pair. We also compare the results of our new approach with Histogram Matching [13].
We measure the differences between the albedo maps computing the Median of the Euclidean Distance between pixels. In the case of shading maps we subtract the mean and normalize the values before computing the difference. This gives a better indication of the difference in the produced geometry, since the shape from shading algorithm assumes a globally flat surface, meaning a zero average.
Figure 11 shows the results of these comparisons for 13 pairs of exemplars. The measured error is shown as the percentage of the maximum error possible (100% error is equivalent to compare a white image with a black image).
In general, the average error is similar for both Histogram Matching (HM) and Transfer by Analogy (TBA). Transferred albedo maps give rise to a mean error of around 3%. Shading produced higher errors (8%), but both methods provide almost identical average performance. Transfer by Analogy also performs a more accurate separation between albedo and shading when the feature is present in the exemplar, as well as being less sensitive to global variations of the texture due to its local matching. This validates our method as a suitable alternative that provides very interesting features that make it more desirable: super-resolution and material segmentation.

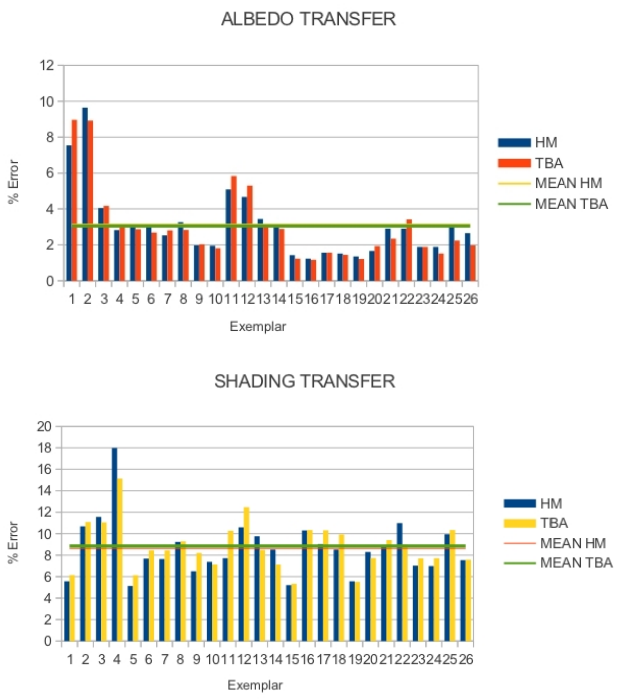
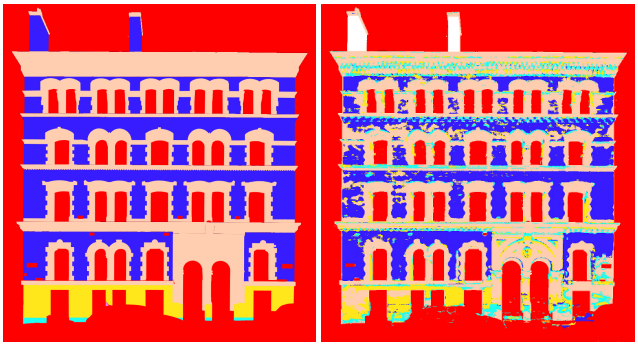


Figure 11: Error comparison between our new transfer method, Transfer by Analogy (TBA), and Histogram Matching (HM).



(a) Manual segmentation using image processing software

(b) Automatic segmentation using Transfer by Analogy

Figure 12: Comparison between the segmentation and association performed by Transfer by Analogy and a reference segmentation created manually by a user.

Figure 13: The lighting conditions in the photo-texture and the exemplar are matched using histogram matching. In the red rectangle, a section of the facade is shown before and after being matched to the exemplar in the blue rectangle.

5. TRANSFER TO LARGE SURFACES

Thus far in this paper, we have shown our transfer techniques applied for only one material and to regions of a similar scale. In order to apply our method to large surfaces, a number of issues need to be solved, most notably the presence of several materials in the photo-texture for a façade (as shown in Figure 15) and the difference of scale and lighting conditions between the exemplar and the photo-texture.

Transfer by Analogy, naturally facilitates searching for the best match between several materials, and therefore, the real strength of the approach is that no segmentation is required for the transfer process. In Figure 12, we show the association of materials provided by Transfer by Analogy by taking the best match among a set of exemplars, compared with the user specified segmentation and association. We can see that visually the result compares well to the ground truth. Some innacuracies appear in the association of similar materials, especially where the texture features are not present in the exemplar, or the association is not clear, like in cornices and ornamentation at the portico. Our Log Based Metric shows good spatial coherence in the associations but still shows some fragmentation (see Figure 12). The segmentation could be further improved by using a regularization term on a *Markov Random Field* formulation, such as Graph-cuts [5].

Transfer by Analogy requires patches in the ambient capture of the exemplar and the photo-texture to be directly compared, therefore it is not invariant to scale and lighting variations. To solve the scale issue we adapt the scale manually by measuring a repeat feature in the photo-texture (for example a brick) and finding the ratio between it and the exemplar. If the exemplar is contained in the photo-texture, as happens in the case shown in Figure 13, this process can be automated using scale-invariant feature detectors [6]. To normalize the lighting conditions between the exemplar ambient image and the photo-texture we select an area in the photo-texture of a similar size to the exemplar and match the histogram of the ambient capture to this section. This effectively adapts the appearance of the exemplar to the lighting conditions of the photo-texture as seen in Figure 13.

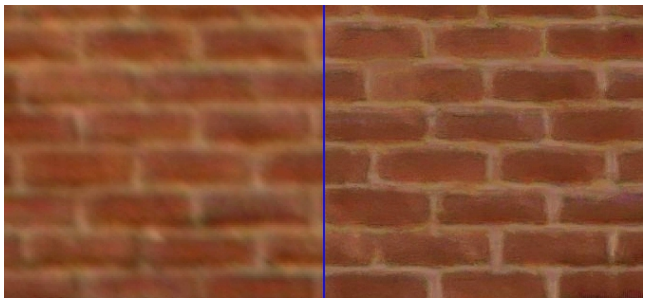


Figure 14: Example of enhancement $2x$ of the texture using Transfer by Analogy.

A particular strength of the Transfer by Analogy process is that the matching finds coordinates in an exemplar which we are able to exploit in order to boost resolution of the photo-texture. Normally exemplars will have a greater resolution since they are typically captured at close range and represent about $1m^2$ of surface while a photo-texture of a building façade would normally be constructed from reference images taken from much further away. Figure 14

shows the effect of using an exemplar with twice the resolution than the original sample (right image). The process to increase the resolution is simple. We scale up the coordinate map that associates the photo-texture to the exemplar and use pixel voting, as described before, to avoid blocky effects.

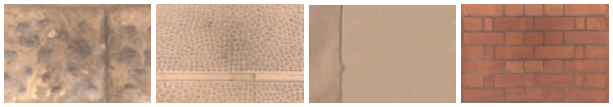
6. RESULTS

In this section we demonstrate the usability of the technique presented in this paper with two case studies for acquisition of albedo and displacement maps for complete façades of real buildings. We capture a high resolution photo-texture for each full façade.

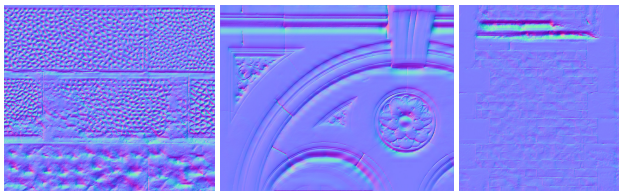
(a) Photo-texture of the Façade (b) Transferred Albedo Map

Figure 15: Captured photo-texture and inferred albedo from the exemplars.

(a) Close up detail of the albedo map transferred with TbA



(b) Exemplar albedo maps for comparison



(c) Close up detail of the normal map (TbA)

Figure 16: Close up views of Transferred maps using Histogram Transfer. Albedo is consistent with the albedo maps captured for the exemplars. Geometric detail preserves the local detail consistently with the appearance of the façade.

Material characteristics are transferred to the photo-texture, automatically generating albedo (shown in Figure 15) and a shading map from which we compute a displacement map. This can be used to further compute a normal map as shown in Figure 16. We show close-up views of the resulting albedo and normal maps, together with the albedo maps of the exemplars for comparison. The overall appearance of the albedo maps as well as the normal maps is compelling and maintains a good degree of detail.

The resulting albedo maps produced are similar and consistent with the albedo of the exemplars. Regarding the geometry created with our method, an interesting result is the case of the ornamentation in the portico (middle column of Figure 16). Our method is able to provide nicely detailed models of the floral motifs even though these features are not present in the exemplar (third exemplar from the left in Figure 16(b)). Figure 18 shows relit models produced by meshing the depth maps created from the transferred shading maps.

Figure 18: Close up renders of meshed results using simple point light sources.

The second case is presented in Figures 17. Here, we also show how the resulting albedo and displacement maps can be applied to a simple polygonal 3D model and generate renderings under novel lighting. We focus on the brick texture, showing how the transferred geometry reproduces self-shadowing effects and the appearance captured in the exemplar.

Although our approach produces good results for a variety of textures, some limitations arise when the original phototextures present hard shadowing, variations that are not present in the exemplars, or occlusions and cannot be matched. In these cases when the phototexture-exemplar match present a large error, other tecniques such as texture inpainting could be used to fill the unmatched areas.

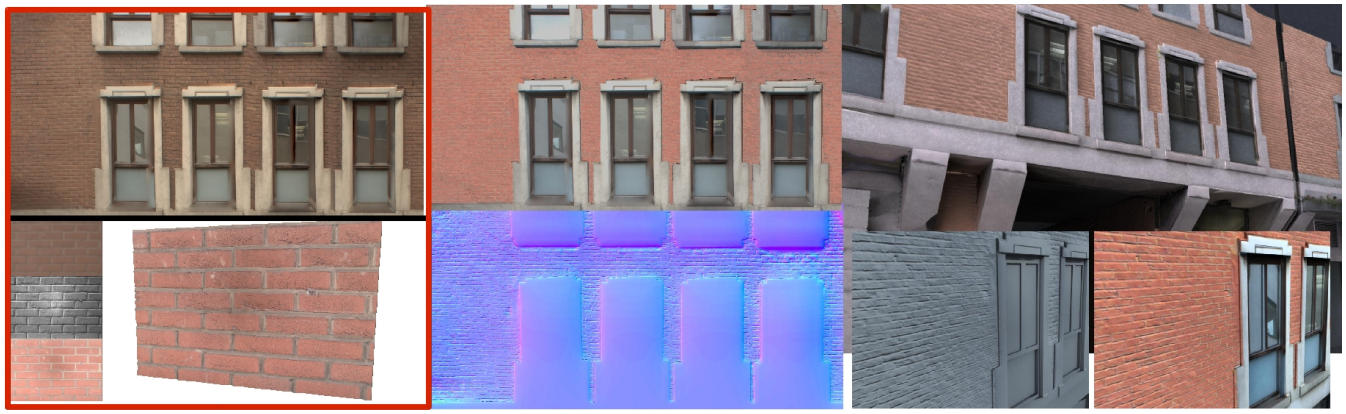


Figure 17: Photo-texture and exemplar of the predominant material (red rectangle), albedo and texture geometry (middle column), enhanced relit 3D model (right).

7. CONCLUSIONS AND FUTURE WORK

In this paper we have presented a solution for fast semi-automatic creation of albedo and displacement maps for full building façades. We detail our new Transfer by Analogy technique together with optimizations for superresolution. Our method provides two important advantages over previous work. First, by performing more localized matching, the technique is less sensitive to the quality of the match between the statistics of the exemplar and the target image. This is important since exemplars are small samples of the material and might contain features that are statistically significant in the exemplar but not in the global texture, or vice versa. In relation to this, we can search for the best patch in several exemplars sidestepping the need for segmenting a texture and fixing a material for every segment. This is a very valuable advantage since texture segmentation can be very time consuming when photo-textures contain many materials. Second, our method facilitates boosting the resolution of the photo-texture captured at a distance with information captured at close range.

As future work, we consider that further investigation of transfer-oriented texture descriptors is a promising field in order to improve the performance of Transfer by Analogy. Multi-resolution descriptors [27] as well as search in multiple orientations [3] are being researched and our techniques could benefit from these. Finally, under diffuse lighting conditions, large façades often present shadowing due to different architectural features. We do not deal explicitly with this, but soft shadow removal from textured surfaces is indeed an interesting research problem.

Finally, using a simple acquisition process and consumer SLR equipment, we demonstrate the efficacy of our processes to significantly enhance the detail recovered for building façades within a full image-based reconstruction and relighting pipeline.

8. ACKNOWLEDGMENTS

The authors wish to than Elliot Newman for providing models and textures to evaluated our technique. Funded was provided by the UK Engineering and Physical Sciences Research Council Knowledge Transfer Account at Loughbourgh University.

9. REFERENCES

[1] M. Ashikhmin. Synthesizing natural textures. In *Symposium on Interactive 3D graphics I3D*, pages 217–226. ACM, 2001.

[2] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman. Patchmatch: a randomized correspondence algorithm for structural image editing. In *SIGGRAPH*, pages 24:1–24:11. ACM, 2009.

[3] C. Barnes, E. Shechtman, D. B. Goldman, and A. Finkelstein. The generalized PatchMatch correspondence algorithm. In *European Conference on Computer Vision*, Sept. 2010.

[4] S. Boivin and A. Gagalowicz. Image-based rendering of diffuse, specular and glossy surfaces from a single image. In *SIGGRAPH*, pages 107–116, Los Angeles, California, 2001. ACM.

[5] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(11):1222–1239, 2001.

[6] M. Brown, R. Szeliski, and S. Winder. Multi-image matching using multi-scale oriented patches. San Diego, June 2005.

[7] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink. Reflectance and texture of real-world surfaces. *ACM Transactions on Graphics (TOG)*, 18(1):1–34, 1999.

[8] P. Debevec, C. Tchou, A. Gardner, T. Hawkins, C. Poullis, J. Stumpfel, A. Jones, N. Yun, P. Einarsson, T. Lundgren, M. Fajardo, and P. Martinez. Estimating surface reflectance properties of a complex scene under captured natural illumination. Technical Report ICT-TR-06.2004, University of Southern California Institute for Creative Technologies, 2004.

[9] P. E. Debevec. Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *SIGGRAPH*, pages 189–198, Orlando, Florida, 1998. ACM.

[10] J. Dorsey, H. Rushmeier, and F. Sillion. *Digital Modeling of Material Appearance*. The Morgan Kaufmann Series in Computer Graphics. Elsevier Science & Technology, 2007.

[11] A. A. Efros and W. T. Freeman. Image quilting for texture synthesis and transfer. In *SIGGRAPH*, pages 341–346. ACM, 2001.

[12] A. S. Georghiades. Recovering 3-D shape and reflectance from a small number of photographs. In *Workshop on Rendering*, pages 230–240, Leuven, Belgium, 2003.

Eurographics.

[13] M. Glencross, G. J. Ward, C. Jay, J. Liu, F. Melendez, and R. Hubbold. A perceptually validated model for surface depth hallucination. *ACM SIGGRAPH*, 27(3):59:1 – 59:8, 2008.

[14] D. J. Heeger and J. R. Bergen. Pyramid-based texture analysis/synthesis. In *SIGGRAPH*, pages 229–238. ACM, 1995.

[15] A. Hertzmann, C. E. Jacobs, N. Oliver, B. Curless, and D. H. Salesin. Image analogies. In *SIGGRAPH*, pages 327–340. ACM, 2001.

[16] B. K. P. Horn. Obtaining shape from shading information. In *Series of Artificial Intelligence: Shape from Shading*, pages 123–171. Mit Press, Cambridge, MA, 1989.

[17] R. M. Ismert, K. Bala, and D. P. Greenberg. Detail synthesis for image-based texturing. In *I3D '03: Proceedings of the 2003 symposium on Interactive 3D graphics*, pages 171–175, New York, NY, USA, 2003. ACM.

[18] J. Kopf, C.-W. Fu, D. Cohen-Or, O. Deussen, D. Lischinski, and T.-T. Wong. Solid texture synthesis from 2d exemplars. *ACM Trans. Graph.*, 26, July 2007.

[19] N. Kumar, L. Zhang, and S. Nayar. What is a good nearest neighbors algorithm for finding similar patches in images? In *Proceedings of the 10th European Conference on Computer Vision: Part II*, pages 364–378, Berlin, Heidelberg, 2008. Springer-Verlag.

[20] V. Kwatra and L.-Y. Wei. Course 15: Example-based texture synthesis. ACM SIGGRAPH Courses Program, 2007.

[21] M. S. Langer and S. W. Zucker. Shape-from-shading on a cloudy day. *Journal of the Optical Society of America*, 11(2):467–478, 1994.

[22] S. Lefebvre and H. Hoppe. Appearance-space texture synthesis. In *SIGGRAPH*, pages 541–548. ACM, 2006.

[23] H. P. A. Lensch, J. Kautz, M. Goesele, W. Heidrich, and H.-P. Seidel. Image-based reconstruction of spatial appearance and geometric detail. *ACM Transactions on Graphics (TOG)*, 22(2):234–257, 2003.

[24] S. R. Marschner, S. H. Westin, E. P. F. Lafortune, and K. E. Torrance. Image-based bidirectional reflectance distribution function measurement. *Applied Optics*, 39(16):2592–2600, 2000.

[25] F. Melendez, M. Glencross, G. J. Ward, and R. Hubbold. Relightable buildings from images. In *Eurographics: Special Area on Cultural Heritage*. Eurographics, 2011.

[26] J. A. Paterson, D. Claus, and A. W. Fitzgibbon. BRDF and geometry capture from extended inhomogeneous samples using flash photography. In *Computer Graphics Forum*, volume 24, pages 383–391. Eurographics, 2005.

[27] E. Risser, C. Han, R. Dahyot, and E. Grinspun. Synthesizing structured image hybrids. In *ACM SIGGRAPH 2010 papers*, SIGGRAPH '10, pages 85:1–85:6, New York, NY, USA, 2010. ACM.

[28] H. Rushmeier and F. Bernardini. Computing consistent normals and colors from photometric data. In *Second Conference on 3-D Imaging and Modeling 3DIM*, pages 99–108, Ottawa, Canada, 1999. IEEE.

[29] L. Shen and H. Takemura. Spatial reflectance recovery under complex illumination from sparse images. In *Computer Vision and Pattern Recognition*, pages 1833–1838. IEEE, 2006.

[30] D. Simakov, Y. Caspi, E. Shechtman, and M. Irani. Summarizing visual data using bidirectional similarity. In *Computer Vision and Pattern Recognition CVPR*, pages 1–8. IEEE, 2008.

[31] G. J. Ward. Measuring and modeling anisotropic reflection. In *SIGGRAPH*, pages 265–272. ACM, 1992.

[32] G. J. Ward and M. Glencross. A case study evaluation: Perceptually accurate textured surface models. In *Symposium on Applied Perception in Graphics and Visualization (APGV)*, pages 109–115. ACM, 2009.

[33] L.-Y. Wei, S. Lefebvre, V. Kwatra, and G. Turk. State of the art in example-based texture synthesis. Technical Report EG-STAR, Eurographics State of the Art Report, 2009.

[34] L.-Y. Wei and M. Levoy. Fast texture synthesis using tree-structured vector quantization. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '00, pages 479–488, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[35] Y. Wexler, E. Shechtman, and M. Irani. Space-time completion of video. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29:463–476, March 2007.

[36] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980.

[37] C. Xu, A. Georghiades, H. Rushmeier, and J. Dorsey. A system for reconstructing integrated texture maps for large structures. In *3DPVT '06: Proceedings of the Third International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'06)*, pages 822–829, Washington, DC, USA, 2006. IEEE Computer Society.

[38] Y. Yu and J. Malik. Recovering photometric properties of architectural scenes from photographs. In *SIGGRAPH*, pages 207–217. ACM, 1998.

[39] R. Zhang, P.-S. Tsai, J. E. Cryer, and M. Shah. Shape from shading: a survey. *IEEE Transactions on Pattern Analysis and MAchine Intelligence*, 21(8):690–706, 1999.