Vision, Modeling, and Visualization (2011)
Peter Eisert, Konrad Polthier, and Joachim Hornegger (Eds.)

Warp-based Near-Regular Texture Analysis for Image-based Texture Overlay

Anna Hilsmann[1,2], David C. Schneider[1,2] and Peter Eisert[1,2]

[1]Fraunhofer Heinrich Hertz Institute, Berlin, Germany
[2]Humboldt University of Berlin, Germany

Abstract

Image-based texture overlay *or* retexturing *is the process of augmenting a surface in an image or a video sequence with a new, synthetic texture. Some properties of the original texture such as texture distortion as well as lighting conditions should be preserved for a realistic appearance of the augmented result. One approach would be to estimate a 3-dimensional geometry of the surface. However, this is an ill-posed problem for complex deformed surfaces like cloth, especially if only one image is given. In an image-based approach, these properties are directly estimated from the image. The key challenge is to separate the shading information from the actual local texture and to retrieve the texture distortion from an image without any knowledge of the underlying scene.*
In this paper, we model an image of a deformed regular texture as a combination of its deformed surface albedo, a shading map and additional high frequency details. We present a method for determination of these intrinsic *parts of a given texture image by first estimating the appearance of a small texture element and then synthesizing a reference image of the undeformed regular texture. In a subsequent image-based optimization method this reference image is iteratively warped spatially and photometrically onto the original image whilst estimating deformation and illumination parameters. The decomposition is used to create images of new textures with the same deformation and illumination properties as in the original image*

Categories and Subject Descriptors (according to ACM CCS): I.4.7 [Image Processing and Computer Vision]: Feature Measurement—Texture I.3.3 [Computer Graphics]: Picture, Image Generation—

1. Introduction

Augmenting a surface in an image with a new synthetic texture (see figure 1) is a challenging problem which has been addressed in recent years both by the Computer Vision and Graphics communities. On the one hand, texture distortion caused by projecting the surface into the image plane should be preserved. On the other hand, only the texture albedo should be altered but shading and reflection properties should remain as in the original image. In many applications, such as augmented reality applications for virtual clothing, the surface material to be retextured is cloth. In this case, high frequency details, representing e.g. self-shadowing of the yarn structure, might also be a property that should be preserved in the augmented result.

This paper specifically addresses the decomposition of images of deformed regular textures into its *intrinsic* parts

Figure 1: Retexturing *means the replacement of texture under preservation of texture deformation and illumination properties.*

which decompose the image into the appearance of the undeformed regular texture, a deformation field, a shading map representing lighting effects and additional high frequency details (see figure 3). This is closely related to intrinsic im-

DOI: 10.2312/PE/VMV/VMV11/073-080

Figure 2: *Many cloth patterns are of a regular type and can be constructed by tiling the space with the same texture element.*

age decomposition methods which decompose an arbitrary image into the product of an illumination component that represents lighting effects and a reflectance component related to the color of the observed material [BPD09, TFA05]. In our decomposition model, the deformed regular texture (the estimated deformation field applied to the estimated appearance of the regular texture) can be seen as the reflectance part of our decomposition while the shading map and the high frequency details can be seen as the illumination component.

Regular textures can be constructed by regularly tiling the texture space with the same texture element, called *texel* in the following [GS86]. Textures that deviate geometrically and photometrically from a regular congruent tiling are often called *near-regular textures* (NRT) [LLH04, LL03]. In contrast to regular textures, the texture elements appear geometrically and photometrically distorted in the image due to variations in the viewing angle, lighting conditions and partial occlusions (see figure 2). Nevertheless, they still exhibit certain topological regularities and relations as regular textures [LLH04, PBCL09]. We exploit this topological regularity to estimate the intrinsic parts of the given image of a near-regular texture. We first start by estimating the mean appearance of a texel and candidate positions of the texel in the image. From the estimated mean texel we synthetically generate an image of the regular texture. This image is used as reference in an image-based optimization method that registers two images not only geometrically but also photometrically, yielding a deformation field and a shading map. Finally, as the mean texel does not contain any high-frequency details, these can be estimated from the difference between the original image and the warped and shaded synthetic reference image.

The remainder of this paper is structured as follows. The next section reviews related work before section 3 explains our image decomposition method. Section 4 finally presents how the proposed method can effectively be used for retexturing purposes.

2. Related Work

Various texture overlay methods have been proposed for videos [PLF05, HSE10, SM06]. These methods estimate surface deformation and shading properties in relation to a given reference frame of the undeformed and uniformly lit texture. Other methods for single images use markers or specifically designed cloth as reference [ES05, WF06b]. However, such a reference image is not always available.

In this paper, we address the problem of automatic texture replacement given a single image of a deformed surface without the knowledge of a reference image of the undeformed surface. This problem is related to shape-from-shading and shape-from-texture problems which use shading or texture deformation as strong cues for depth to reconstruct the 3D structure of an object from a single image [WF06a, HSE11]. Consequently, current methods for single-image retexturing either use shading [FH04, GSPJ08, YS10] or texture [LLH04] to estimate the deformation field of the texture and additional lighting conditions. Fang and Hart [FH04] estimated surface normals from shading (edited by the user) and proposed a normal-guided texture synthesis method that produced compelling replacement effects. Similarly, Yan and Shen [YS10] decompose the input image into intrinsic images and use the reflectance component to estimate the surface normals which guide the synthesis approach. These methods are limited to untextured, diffuse surfaces illuminated by a single directional light source.

Liu et al. [LLH04] presented an interactive method for near-regular texture analysis and manipulation. They categorize near-regular textures based on their geometric and color irregularities (based on their categorization the type of texture this paper addresses would be NRT III - irregular geometry and irregular color). They introduced an underlying topological lattice structure for NRT. The basic idea is that for each near-regular texture there is a lattice that describes a deformation of that texture from a regular lattice. The lattice generation needs interactive and high accurate editing by the user. Once the geometric deformation is known, the texture is straightened out and a light map is extracted using a method from Tsin et al. [TLR01]. This procedure is similar to intrinsic image estimation methods that model an image as a product of an illumination component representing shading and illumination and a reflectance component representing the color of the observed material [BPD09, TFA05].

Recently, automatic lattice detection methods have been proposed for 2D wallpapers in real-world images to overcome the user interaction during lattice generation

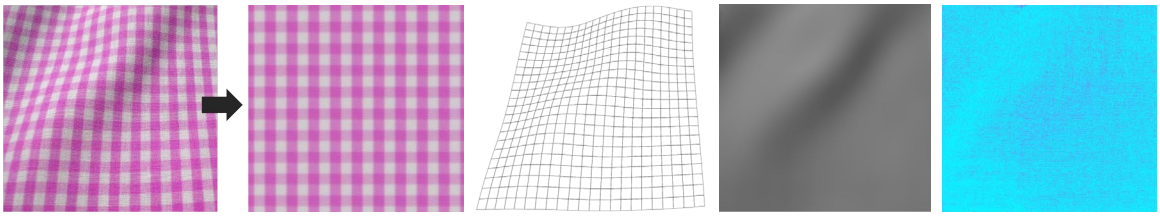

Figure 3: *Decomposition of the original texture (left) according to equation (1) into a regular texture $\mathcal{T}(\mathbf{x})$, deformation field $\mathcal{W}_g(\mathbf{x})$ represented by a deformed mesh, estimated shading map $\mathcal{W}_p(\mathbf{x})$, estimated high frequency yarn structure $\mathcal{HF}(\mathbf{x})$ (colors scaled).*

[HLEL06, PBCL09]. Park et al. [PBCL09] use unsupervised clustering of interest points to detect repeating elements in the image and exploit an assumed topology of wallpaper patterns to initialize a pair of basis vectors for lattice generation. The mesh is refined and grown in a spatial multitarget tracking problem solved in a Mean-Shift Belief Propagation method.

In our approach, we decompose the input image of a deformed regular texture into its intrinsic parts, i.e. components that represent the surface albedo or reflectance and illumination components that represent shading effects. The reflectance component is furthermore decomposed into the image of an undeformed texture and a deformation field. The shading component consists of a global shading map and additional high frequency details that represent stochastic color irregularities, e.g. due to self shadowing at the yarns in cloth. We represent the deformation field similarly to Liu et al. [LLH04] as a deformed mesh describing the deformation of the texture from a regular one. In contrast to [LLH04], we estimate the deformation not only automatically but also simultaneously with a photometric warp in an image-based optimization approach. We exploit the assumption that the original texture is of a regular type, which is the case for many cloth textures (e.g. figure 2). The key idea is to separate the surface albedo from photometric or shading information by estimating the mean appearance of one texture element and to estimate the texture deformation and shading as a spatial and photometrical deviation from a regular texture synthesized from the estimated texel appearance. Our method for texel appearance estimation is inspired by the lattice unit detection approach of [PBCL09] and based on mean-shift clustering of feature points in the image. Finally, the decomposed parts are combined with any new texture to generate realistic retexturing results. The composition of high frequency shading effects and yarn structure with the new texture can be manipulated by a user.

3. Texture Analysis and Decomposition

In our approach, we model a given image $\mathcal{I}$ of the deformed regular texture as

$$\mathcal{I}(\mathbf{x}) = \mathcal{W}_p(\mathbf{x}) \cdot \mathcal{T}(\mathcal{W}_g(\mathbf{x})) + \alpha \cdot \mathcal{HF}(\mathbf{x}) \qquad (1)$$

where $\mathbf{x}$ denotes pixel coordinates, $\mathcal{T}$ is the original regular texture which is deformed by a geometric warp $\mathcal{W}_g(\mathbf{x})$, $\mathcal{W}_p(\mathbf{x})$ denotes a photometric warp, i.e. a shading map multiplied to the image intensities, and $\mathcal{HF}(\mathbf{x})$ denotes high frequency details on the texture. Figure 3 illustrates this image model. The idea behind our approach for texture overlay is to decompose an image of a deformed regular texture into the different components to substitute $\mathcal{T}$ with a new texture. For this purpose, we exploit the assumption that the original texture $\mathcal{T}$ is regular, i.e. it can be constructed by regularly tiling the texture space with the same texel (note that there are more than one valid texels as each shifted version of a valid texel or a set of more than one valid texel are also valid texels). Thus, the deformed texture $\mathcal{T}(\mathcal{W}_g(\mathbf{x}))$ also exhibits repeated similar elements (up to a geometric and photometric deformation) with the same topological relations. The proposed method consists of the following substeps:

- **Mean texel appearance and lattice estimation**.
  In a first step we estimate the mean appearance of one repeating texture element and a lattice structure representing the topological relations between texels and candidate texel positions in the image. From the mean texel appearance an estimated appearance of the regular texture $\mathcal{T}$ of arbitrary size can be synthesized (section 3.1).

- **Texture decomposition through joint geometric and photometric registration**.
  In an image-based optimization procedure we jointly estimate the geometric deformation $\mathcal{W}_g(\mathbf{x})$ and the shading map $\mathcal{W}_p(\mathbf{x})$ by registering $\mathcal{T}$ onto $\mathcal{I}$ geometrically and photometrically. As the synthetic regular texture is generated from an estimated mean texel, high frequency parts of the original texture, the remaining residual between the original image and the registered synthetic texture represents an estimate of these structures (section 3.2).

- **Texture replacement.**
  Having decomposed the image into its intrinsic parts, a new synthetic texture (not necessarily a regular one) can now be substituted into equation (1) to produce an image of this texture with the same deformation and

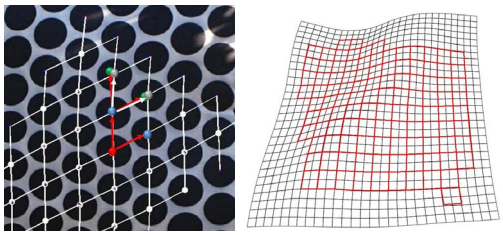

Figure 4: *Left: Illustration of lattice generation from feature points. Starting at a seed point (red) and two proposing vectors, two new lattice vertices are found (blue). These points inherit the* proposing *vectors, new lattice positions are proposed (green) and cluster points near the proposed positions are taken as lattice vertices. The* proposing *vectors are updated with the correct vectors between the points (white). The same procedure is done for the negative proposing vectors. Right: The mesh (black) we use for registration is finer than the lattice (red) and the lattice positions are used to initialize the mesh.*

illumination properties (section 3.3).

3.1. Mean Texel Appearance and Lattice Estimation

We start by first estimating the mean appearance of one or more texels and candidate positions of these texels in the image, represented by a quadrilateral lattice. Regular textures can be described by one texture element and two smallest linear independent generating vectors [GS86]. In photometrically and geometrically deformed regular patterns, the texel appearances in the image are no longer copies of each other but rather geometrically deformed and of different colors due to shading. Nevertheless, the appearances of the texels are still *similar* (e.g. up to geometric and photometric distortion as in our model). To identify the texel appearances in the image, we follow a similar strategy as [PBCL09]. We start by generating suitable feature points on the image and group the descriptors using mean shift clustering [CM02] to identify repeating structures in the image. Although the texture in the image may be strongly distorted, the idea is that transformations between local image patches can be approximated by a projective, an affine or even a similarity transformation (this assumption is similar to the planarity assumption used in most shape from texture methods [WF06a]). Hence, ideally a feature descriptor would be invariant against small projective transformations. Park et al. [PBCL09] use normalized image patches of a fixed size as descriptors which are only invariant against translation such that strongly rotated or scaled image patches cannot be detected. To this end we use SIFT features [Low04] in our approach. Although this descriptor is only invariant against rotation and scale and not against affine or perspective distortions which appear at very strong deformations, it produces good detection results if there are not too many strongly distorted texels.

Each cluster $\mathcal{C}_i : \{\mathbf{p}_j, j = 1 \dots n\}$ now consists of a set of n image points $\mathbf{p}_j$ with similar SIFT descriptors. The image points of one cluster should by related to each other by the same topological regularities and relations as the undeformed regular texture. According to [LLH04], [PBCL09] any deformed regular texture can be described by a degree-4-graph representing the tiling pattern of the texels in the undeformed texture. Note that as there are more than one valid texels there are also more than one valid tiling patterns. The aim of this step is to estimate the appearance of the repeating texture element and a quadrilateral lattice model L_i for each cluster, consistent with the geometric relationship between the feature points and the assumed texture topology. In this lattice, the texture lying in each quad approximates a texel appearance $t_i^k, k = 1 \dots m$ coarsely deformed by a homography defined by the four quad vertices. In the following, we use the term *quad* as a topological element of the lattice, consisting of 4 vertices, and *texel* as the image part lying between those four vertices. Although the real deformation of the texels might be more complex, we use the lattice to estimate a mean appearance of the texels and refine the deformation in a subsequent image-based registration step (section 3.2).

To estimate the lattice structure, we start with a seed point $\mathbf{p}_S$ in the center of all cluster image points. For this seed point, we search for the nearest 8 neighbors and define an L-shaped vector pair $\mathbf{v}_1$, $\mathbf{v}_2$ pointing to two of its neighbors. From all possible vector pairs from the seed to the eight neighbors, we choose the one with the most perpendicular angle. Now, we search for four nearest cluster points that are in a predefined distance (dependent on the mean distance between cluster points) from $\mathbf{p}_S \pm \alpha\mathbf{v}_1$ and $\mathbf{p}_S \pm \alpha\mathbf{v}_2$, $0.5 \leq \alpha \leq 2$. If such points are found, they become vertices of the lattice and the lattice edges are defined by the vectors from the seed to these points. For each of these four points we update the *proposing* vectors $\mathbf{v}_1, \mathbf{v}_2$ and $-\mathbf{v}_1, -\mathbf{v}_2$ by the real vectors pointing from the seed to its children and we proceed for all children as we did for the seed point. From the found edges between cluster points we generate a quadrilateral lattice with candidate quads which are evaluated as explained in the following. The lattice detection procedure is schematically illustrated in figure 4.

To reject wrongly detected quads, we rectify each texel associated to the quad candidates into a normalized texel coordinate system and normalize it by subtracting the mean and dividing through the standard deviation. From these normalized texel candidates, a rectified and normalized mean texel appearance is calculated. A similarity measure (e.g. the sum of squared differences of pixel values) is calculated between each texel and this mean. This similarity measure is used to reject wrongly detected texels and quads with a MAD-based (median of absolute differences) outlier rejection method [IH93]. Note that the rectification to a normalized texel coordinate system is done regardless of the true

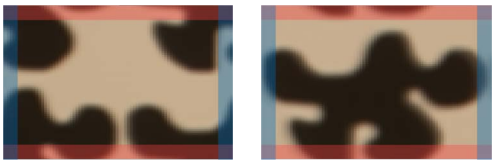

Figure 5: *Each valid rectified and normalized texel should be self-similar at the upper and lower (red) and at the right and left border (blue) respectively.*

texel shape to compare the appearances of the texel quads and to estimate a mean rectified appearance.

Additionally, we validate each remaining texel based on a continuity measure. A valid rectified texel of a regular texture should have similar upper and lower as well as left and right borders to produce seamless transitions during tiling (see figure 5). Hence, the continuity measure for each texel is based on the sum of squared difference of pixel intensities of the left and right borders and the upper and lower borders, respectively. We reject quads with a bad continuity measure with a MAD-based outlier rejection method as before. From the remaining quads we construct a topologically consistent quadrilateral lattice L_i and calculate a mean texel appearance $\bar{t}_i$.

To conclude, for each cluster $\mathcal{C}_i$ we have estimated a mean rectified texel appearance $\bar{t}_i$ and a lattice structure L_i. For readability reasons we will skip the index $_i$ for the cluster in the following.

3.2. Texture Decomposition through Joint Geometric and Photometric Registration

In the previous step we estimated a mean rectified appearance of one valid texel $\bar{t}$ and coarse positions of texel appearances in the image represented by a lattice L. By exploiting the assumed topology of the texture regularity, we synthesize an image $\mathcal{T}$ of the undeformed texture from the estimated mean texel $\bar{t}$ by regular tiling. Thereby, we synthetically generate a reference image of the undeformed texture. The estimation of the texture deformation and the shading map is now treated as a geometric and photometric image registration task, solving for a warp that registers the synthesized undeformed texture onto the original image not only geometrically but also photometrically.

We jointly estimate a deformation and a shading map (see figure 3) using an image-based optimization scheme (similar to [HSE10]). This optimization scheme starts from a relaxed brightness constancy equation and formulates a pixel-wise error at pixel $\mathbf{x}_i$ as

$$r_i(\theta) = \mathcal{W}_p(\mathbf{x}_i;\theta_p)\cdot\mathcal{T}(\mathcal{W}_g(\mathbf{x}_i;\theta_g)) - \mathcal{I}(\mathbf{x}_i) \qquad (2)$$

$\mathcal{W}_p$ and $\mathcal{W}_g$ are photometric and geometric warp functions parameterized by photometric and geometric parameters $\theta = [\theta_g^T\,\theta_p^T]^T$.

Both warps are mesh-based. Note that we differentiate between the *mesh M* which we use as motion model to register the two images and the *lattice L* describing the texture topology and coarse texel positions in the image. Let M_r denote a regular undeformed mesh on $\mathcal{T}$ and $\mathbf{V}^{M_r} : \{\mathbf{v}_1^{M_r} \dots \mathbf{v}_K^{M_r}\}$ denote its set of vertices. The mesh has a finer structure than the lattice to allow for more complex deformations of the texel. Its topology is a rectangular grid with additional diagonal edges through the quads. If L_r denotes the regular lattice on $\mathcal{T}$ with vertices $\mathbf{V}^{L_r} : \{\mathbf{v}_1^{L_r} \dots \mathbf{v}_H^{L_r}\}$ corresponding to the deformed lattice L we choose M_r such that $\mathbf{V}^{L_r} \subset \mathbf{V}^{M_r}$ and initialize the vertex positions of M by warping M_r onto the lattice L using thin-plate spline (TPS) warping with the lattice vertices as control points (see figure 4). We iteratively grow the mesh (and with it the reference regular texture) after each optimization has reached its minimum and start a new optimization with the grown mesh. In each iteration new mesh vertices are initialized using TPS warping as before.

The geometric warp $\mathcal{W}_g$ of $\mathcal{T}$ onto $\mathcal{I}$ deduced by the mesh is parameterized by vertex displacements. The photometric warp $\mathcal{W}_p$ is parameterized by additional intensity scaling parameters ρ at each vertex. Hence, the geometric parameter vector θ_g contains the concatenated vertex displacements in x- and y-direction and the photometric parameter vector θ_p contains the intensity scaling parameters ρ of all vertices:

$$\begin{aligned}\theta_g &= [d_{1x}\dots d_{Kx}, d_{1y}\dots d_{Ky}]^T \\ \theta_p &= [\rho_1\dots\rho_K]^T\end{aligned} \qquad (3)$$

The warps are then defined as

$$\begin{aligned}\mathcal{W}_g(\mathbf{x}_i;\theta_g) &= \mathbf{x}_i + \mathbf{B}_g^i\cdot\theta_g \\ \mathcal{W}_p(\mathbf{x}_i;\theta_p) &= \mathbf{B}_p^i\cdot\theta_p\end{aligned} \qquad (4)$$

where $\mathbf{B}_g^i$ and $\mathbf{B}_p^i$ depend on the pixel position:

$$\begin{aligned}\mathbf{B}_g^i &= \begin{bmatrix}\mathbf{b}^i & \mathbf{0}\\ \mathbf{0} & \mathbf{b}^i\end{bmatrix} \\ \mathbf{B}_p^i &= \mathbf{b}^i\end{aligned} \qquad (5)$$

and $\mathbf{b}^i$ is a row vector containing only 3 non-zero elements β_a,β_b,β_c at positions a,b,c if the pixel $\mathbf{x}_i$ lies in a triangle built by the vertices $\mathbf{v}_a,\mathbf{v}_b,\mathbf{v}_c$ and β_a,β_b,β_c are the Barycentric parameters.

Estimating θ amounts to minimizing a cost function based on the sum of a pixelwise cost function and an additional smoothness term:

$$\hat{\theta} = \arg\min_{\theta}\mathcal{E}_D(\theta) + \lambda^2\mathcal{E}_S(\theta) \qquad (6)$$

with

$$\begin{aligned}\mathcal{E}_D(\theta) &= \sum_i \psi(r_i(\theta)) \\ \mathcal{E}_S(\theta) &= \Gamma\cdot\theta\end{aligned} \qquad (7)$$

where ψ is a robust norm-like function (we use the Huber

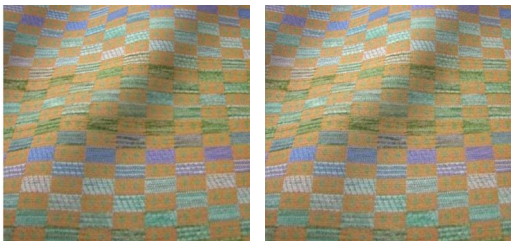

Figure 6: *Retexturing results achieved with $\alpha = 0$ and $\alpha = 1$. Note how the addition of the high frequency details increases the retexturing results.*

function) and Γ is a block diagonal matrix composed of three orientation separated mesh Laplacians, one for each vertex parameter:

$$\Gamma = \begin{bmatrix} \mathbf{L} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \lambda_p \mathbf{L} \end{bmatrix} \tag{8}$$

λ_p weights the smoothing of the photometric parameter against the smoothing of the geometric parameters. The cost function in equation (6) can be minimized in a Gauss-Newton approach that differs only in a weighting scheme from the standard least squares [MN89]. The Gauss-Newton optimization method requires the Jacobians of both the data and the smoothness term. The i^{th} rows of the data term Jacobian $\mathbf{J}_D$ is the gradient of the pixel error r_i:

$$\frac{\partial r_i}{\partial \hat{\theta}} = \mathcal{W}_p(\mathbf{x}_i;\hat{\theta}_p) \cdot \nabla \mathcal{T}(\mathcal{W}_g(\mathbf{x}_i;\hat{\theta}_g)) \cdot \mathbf{B}_g^i + \mathcal{T}(\mathcal{W}_g(\mathbf{x}_i;\hat{\theta}_g)) \cdot \mathbf{B}_p^i \tag{9}$$

where $\nabla \mathcal{T} = [\mathcal{T}_x \mathcal{T}_y]$ is the gradient of the texture image. The smoothness term Jacobian is $\mathbf{J}_S = \Gamma$.

The optimization of θ yields vertex positions of the deformed mesh M and an additional photometric parameter at each vertex position. From the photometric parameters a smooth shading map $\mathcal{W}_p(\mathbf{x}, \hat{\theta}_p)$ is interpolated with bicubic interpolation.

As the synthetic regular texture $\mathcal{T}$ is generated from an estimated mean texel, high frequency parts of the original texture, representing e.g. detailed self-shadows of the yarn structure in cloth, are not present in the synthetic texture. Thus, the remaining residual between the original image $\mathcal{I}$ and the warped and shaded synthetic texture represents an estimate of these structures (see figure 3):

$$\mathcal{HF}(\mathbf{x}) = \mathcal{W}_p(\mathbf{x};\hat{\theta}_p) \cdot \mathcal{T}(\mathcal{W}_g(\mathbf{x};\hat{\theta}_g)) - \mathcal{I}(\mathbf{x}) \tag{10}$$

3.3. Texture Replacement

Having processed and decomposed the input image, equation (1) can now be applied to any arbitrary new synthetic texture to generate an image of this texture with the same deformation and illumination properties as the input image. The weight α can be used to modify the influence of the high detail texture structures.

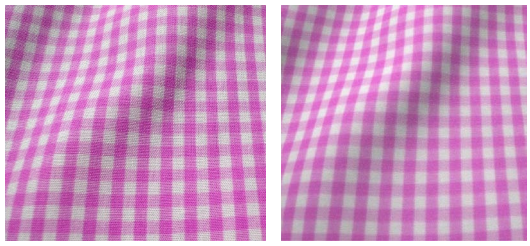

Figure 7: *Left: original image. Right: synthetically generated image accodring to equation (1) with the estimated intrinsic parts as depicted in figure 3.*

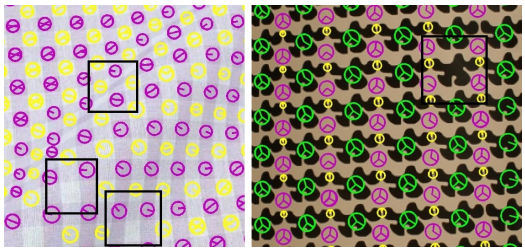

Figure 9: *Missing feature points of two example textures indicated by black squares.*

4. Results

We tested our texture decomposition on several images showing cloth textures. To detect repeating structures in the image we use SIFT feature descriptors. The choice of the descriptor is a trade-off between finding enough feature points to identify repeating structures in the image and creating too many false positive detection results. Although the SIFT descriptor is not invariant against affine or projective transformations (and hence misses some strongly distorted structures, see figure 9) it detected enough feature points to estimate the texel appearance and an initial lattice structure in our experiments. The lattice structure is then refined using image-based optimization. Results for the texel appearance estimation for a variety of regular cloth patterns are depicted in figure 8. Figure10 shows detected feature points, clustering results and the detected initial lattice structure for two examples.

The quality of retexturing results is best evaluated visually. Figure 7 compares a synthetic texture generated from the estimated intrinsic parts to the original image. Although a marginal difference between the images is noticeable when compared directly, the visual appearance of the synthetic result is still very realistic. Figure 6 directly compares retexturing results of the same image achieved with $\alpha = 0$ (left) and $\alpha = 1$ (right). Note how the addition of the high frequency details influences the realistic appearance of the retexturing result. Further retexturing results achieved with $\alpha = 1$ are presented in figure 11 and figure 12 shows details when changing the value for α.

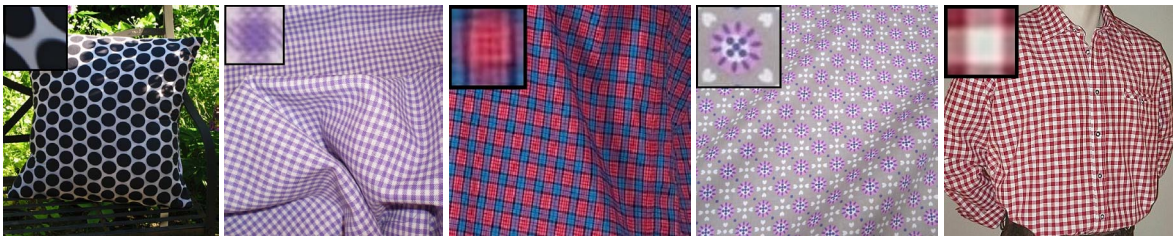

Figure 8: *Different texel appearance estimation results.*

Figure 10: *From left to right: detected feature points, feature clusters marked in different colors and estimated lattice structure for two clusters with the estimated mean texel appearance.*

5. Conclusions and Future Work

We presented an approach for decomposing an image of a deformed regular texture into its intrinsic parts for automatic texture overlay given a single image. We estimate the appearance of a repeating texture element and synthetically generate an image of the undeformed texture. This image is used as a reference image in a subsequent geometric and photometric registration step, yielding a deformation grid and a shading mesh. To detect repeating structures in the image, we currently use the SIFT feature descriptor. However, SIFT fails for strongly foreshortened texels which appear due to perspective distortion. As for the subsequent image-based optimization step a good lattice initialization is needed due to the repetitive texture structure, we will also investigate other (e.g. affine invariant) feature descriptors. Furthermore, self-occlusions due to strong creases and folds lead to discontinuities in the 2D deformation which have not been handled so far. One approach will be to approximate the 3D surface shape from the texture deformation and detect these discontinuities from depth discontinuities.

References

[BPD09] BOUSSEAU A., PARIS S., DURAND F.: User Assisted Intrinsic Images. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia 2009) 28*, 5 (2009). 2

[CM02] COMANICIU D., MEER P.: Mean Shift: A Robust Approach toward Feature Space Analysis. *IEEE Trans. Pattern Analysis and Machine Intelligence 24* (2002), 603–619. 4

[ES05] EHARA J., SAITO H.: Texture Overlay onto Deformable Surface for Virtual Clothing. In *Proc. Int. Conf. Artificial Reality and Telexistence (ICAT 2005)* (2005), pp. 171–178. 2

[FH04] FANG H., HART J.: Textureshop: Texture Synthesis as a Photograph Editing Tool. *ACM Trans. on Graph.* (2004). 2

[GS86] GRÜNBAUM B., SHEPHARD G. C.: *Tilings and Patterns*. W. H. Freeman & Co., New York, NY, USA, 1986. 2, 4

[GSPJ08] GUO Y., SUN H., PENG Q., JIANG Z.: Mesh-Guided Optimized Retexturing for Image and Video. *IEEE Trans. on Visualization and Computer Graphics 14*, 2 (2008), 426–439. 2

[HLEL06] HAYS J. H., LEORDEANU M., EFROS A. A., LIU Y.: Discovering texture regularity as a higher-order correspondence problem. In *Proc. Europ. Conf. on Computer Vision (ECCV 2006)* (Graz, Austria, May 2006). 3

Figure 11: *Different retexturing results achieved with* $\alpha = 1$. *The most left images depict the original input image.*

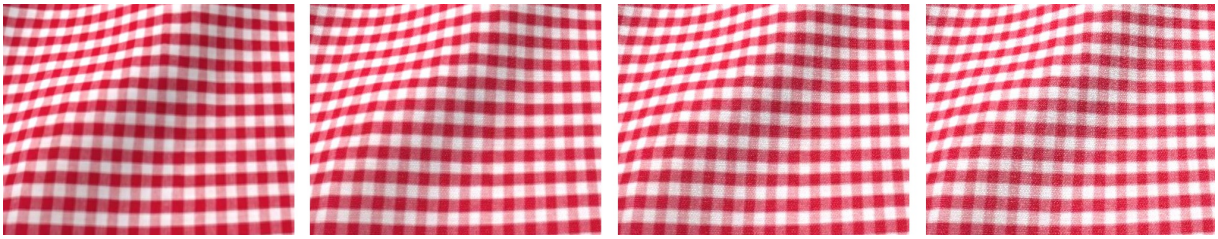

Figure 12: *Influence of different weights for the high frequency structure (left to right* $\alpha = 0, 0.2, 0.6, 1$*).*

[HSE10] HILSMANN A., SCHNEIDER D., EISERT P.: Realistic Cloth Augmentation in Single View under Occlusion. *Computers & Graphics 34*, 5 (2010). 2, 5

[HSE11] HILSMANN A., SCHNEIDER D. C., EISERT P.: Template-free Shape-from-Texture using Perspective Cameras. In *Proc. British Machine Vision Conference (BMVC 2011)* (Dundee, Scotland, 2011). 2

[IH93] IGLEICZ B., HOAGLIN D.: How to Detect and Handle Outliers. *The ASQC Basic References in Quality Control: Statistical Techniques 16* (1993), 10–13. 4

[LL03] LIU Y., LIN W.-C.: *Deformable Texture: the Irregular-Regular-Irregular Cycle, Carnegie Mellon University, 2003*. Tech. rep., 2003. 2

[LLH04] LIU Y., LIN W.-C., HAYS J.: Near-Regular Texture Analysis and Manipulation. *ACM Trans. on Graph.* (2004). 2, 3, 4

[Low04] LOWE D.: Distinctive Image Features from Scale-Invariant Keypoints. *Int. Journ. of Computer Vision 60*, 2 (2004), 91–110. 4

[MN89] MCCULLAGH P., NELDER J. A.: *Generalized Linear Models (Second edition)*. London: Chapman & Hall, 1989. 6

[PBCL09] PARK M., BROCKLEHURST K., COLLINS R., LIU Y.: Deformed Lattice Detection in Real-World Images using Mean-Shift Belief Propagation. *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), Special Issue on Probabilistic Graphical Models 31*, 1 (2009). 2, 3, 4

[PLF05] PILET J., LEPETIT V., FUA P.: Augmenting Deformable Objects in Real-Time. In *Int. Symposium on Mixed and Augmented Reality* (2005). 2

[SM06] SCHOLZ V., MAGNOR M.: Texture Replacement of Garments in Monocular Video Sequences. In *Rendering Techniques 2006: Eurographics Symposium on Rendering* (June 2006), pp. 305–312. 2

[TFA05] TAPPEN M. F., FREEMAN W. T., ADELSON E. H.: Recovering Intrinsic Images from a Single Image. *IEEE Trans.on Pattern Analysis and Machine Intelligence 27* (2005), 1459–1472. 2

[TLR01] TSIN Y., LIU Y., RAMESH V.: Texture Replacement in Real Images. In *Proc. Int. on Computer Vision and Pattern Recognition (CVPR 2001)* (Kauai, HI, USA, 2001), vol. 2, pp. 539 – 544. 2

[WF06a] WHITE R., FORSYTH D. A.: Combining Cues: Shape from Shading and Texture. In *Proc. Int. Conf. on Computer Vision and Pattern Recognition (CVPR 2006)* (New York, NY, USA, 2006). 2, 4

[WF06b] WHITE R., FORSYTH D. A.: Retexturing Single Views Using Texture and Shading. In *Proc. Europ. Conf. on Computer Vision (ECCV 2006)* (Graz, Austria, May 2006), pp. 70–81. 2

[YS10] YAN X., SHEN J.: Mesh-guided Texture Replacement using Intrinsic Images. In *Int. Conf. on Progress in Informatics and Computing (PIC 2010)* (2010), vol. 2, pp. 701 –705. 2