Image Based Proximate Shadow Retargeting

Llogari Casas[1][†], Matthias Fauconneau[1], Maggie Kosek[2], Kieran Mclister[3] and Kenny Mitchell[1]

[1]Edinburgh Napier University, United Kingdom
[2]University of Surrey
[3]Edinburgh College of Art, United Kingdom

a) Ground truth toy lion on textured background   *c) Our uniform image based shadow*

b) Regular virtual projected shadow map   *d) Our image retargeted shadow*

Figure 1: *a) Reference physical toy lion as observed by a mobile phone camera. b) The virtual lion rendered at the same pose with b) a regular shadow map approach and c) with Shadow Retargeting. d) Its movement retargeted according to a sequence of geometry poses in AR towards seamless appearance preservation.*

Abstract
We introduce Shadow Retargeting which maps real shadow appearance to virtual shadows given a corresponding deformation of scene geometry, such that appearance is seamlessly maintained. By performing virtual shadow reconstruction from un-occluded real shadow samples observed in the camera frame, we recover the deformed shadow appearance efficiently. Our method uses geometry priors for the shadow casting object and a planar receiver surface. Inspired by image retargeting approaches [VTP*10] we describe a novel local search strategy, steered by importance based deformed shadow estimation. Results are presented on a range of objects, deformations and illumination conditions in real-time Augmented Reality (AR) on a mobile device. We demonstrate the practical application of the method in generating otherwise laborious in-betweening frames for 3D printed stop motion animation.

CCS Concepts
•*Computing methodologies* → *Mixed / Augmented reality; Graphics systems and interfaces; Perception;*

† Email: l.casascambra@napier.ac.uk

1. Introduction

Recent advances in high performance mobile computing allow enhanced real-time capabilities for Mixed Reality applications. This trend enables our goal to combine the virtual and physical worlds seamlessly together and create a convincing illusion of movement for normally static and inanimate objects. In this work, we adapt image based rendering concepts [KS00] to the novel topic of shadow retargeting in the context of known physical geometry which have been previously scanned or 3D printed.

Typically, image retargeting techniques [VTP*10] perform content-aware 2D resizing through a weighted map that is segmented and aims to retain important features through saliency guided deformations. In our case, with prior scene geometry, we perform retargeting of shadowed imagery with a warping through proximity sample search in the projected 3D shadow. This affords greater flexibility to animate scene content while retaining the importance and visual consistency of features. In such a way, we first make use of a projected 3D mesh registered with physical objects to provide a real-time method for general 3D deformed mesh-based image retargeting (see Fig. 1 for 3D mesh morphed frames). Previous real-time instances of this approach have achieved direct retargeting using only basic cube based deformations [LLT*11] and rectangular building transformations [TP12] and these were further limited to pre-set manual configurations. Of course, directly retargeting surface appearance does not take into account global illumination or view dependent material effects, and without the corresponding shadow appearance these methods are severely limited. Therefore, we focus here on retargeting the appearance of shadows when virtually deforming physical objects in real-time. We map the real shadow to the projected virtual shadow's footprint with an importance guided search according to physically modelled shadow behavior. This forms the weighted reference for appearance sampling when performing reconstruction of the deformed shadow.

Shadow Retargeting also demonstrates the practical application of the method for Stop Motion Animation. It generates otherwise laborious *In-betweening* frames with large reductions in 3D printing costs. Additionally, we avoid the need to process the scene with expensive radiometric calibration, temporal smoothing or complex illumination and material basis re-projections. We highlight the primary contributions as follows:

- we introduce *Proximate Shadow Retargeting*, an efficient and focused method in which already present shadows from static real objects in the scene are retargeted according to virtual overlaid movement over time (Fig. 1) without extensive reflectance and illumination estimation.
- inspired by *Image Retargeting* [VTP*10], we define the goals of shadow retargeting to preserve important shadow content and structure, such that artifacts are minimised whilst exploiting constraints of shadow appearance behavior.
- we describe a discretized concentric ring search algorithm for real-time importance-based shadow sample selection with coherent results.
- we demonstrate shadows retargeted in application to stop motion in-betweening for reduced production costs.

2. Related Work

2.1. Retargeting Reality

Image Retargeting, also known as content-aware resizing, aims to retarget an image to a form with visual constraints (e.g. size) meanwhile preserving its main features. Maintaining an aesthetically plausible output when performing deformations on an image can be a challenging problem to address. Approaches from Setlur *et al.* [STR*05] and Goferman *et al.* [GZMT10], use content-aware importance maps and regions of interest to perceptually guide and steer image sampling in the retargeted reconstruction. Here, we apply the concept of image retargeting to Augmented Reality (AR), in the case where we have known segmented 3D geometry of physical objects in the image and the ability to register the mesh pose in real-time with markerless 3D object tracking [CMPC06] (using *Vuforia* [Vuf17]).

Image retargeting approaches for AR are relatively novel and the first method introduced in 2011 [LLT*11] is a relatively isolated work. Prior to this, virtual objects have been typically overlaid into the physical space as an independent entity or through re-rendering with elaborate light estimation [NGM*11] using a visible chrome sphere. *Altered Reality* [LLT*11] provides a method in which a homogeneous mesh is projected on top of its physical counterpart; a simple deformed cube in this first instance. This demonstration was limited as it neither addressed re-lighting through the direct cube transformation employed nor did the approach address shadows as indirect re-lighting. Additionally, real-time performance was limited to high-end desktop PCs. In contrast, our method introduces direct and indirect relighting on mobile devices in real-time.

In 2012, *Clayvision* introduced a related scheme for building scale objects in the city. This method pre-scanned a set of known locations into reconstructed meshes. The live camera feed was patched with a pre-modified version of the background image, in which the appropriate building was previously removed and manually inpainted. Upon the movement of the building's virtual animated mesh, the desired deformed pixels were overlaid on the camera image [TP12]. While this achieved real-time performance in mobile devices, it did not address any means of indirect relighting.

More recently, real-time texture retargeting from a camera feed has been achieved through the conversion of 2D hand-painted drawings to a 3D colored mesh [MNZ*15]. In this approach, a static lookup map was built in order to match every pixel of the drawing to corresponding sample locations on the mesh. Our method extends this with a dynamic texture lookup according to the overlaid animation over time for the purpose of shadow retargeting.

2.2. Image Inpainting

As defined in the Severely Mediated Virtuality Continuum [Man01], in an ideal Mixed Reality environment, the user should be able to add and subtract visual content on equal basis. Image inpainting, also known as context-aware fill or Diminished Reality (when done in real-time), aims to subtract information from a live camera feed achieving plausible results.

A related case for shadow inpainting is [ZZX15], in which an

illumination recovering optimization method is used for shadow removal. However, our approach focuses on reconstructing it for another pose instead.

Scene reconstruction with structural and perceptual consistency has demonstrated to be computationally expensive over time [BSCB00] [BVSO03]. In 2012, *PixMix* achieved inpainting in real-time with few background constraints and without the need of a multi-view approach [HB12]. This algorithm iteratively minimizes a spatial and appearance cost function, with consistency of neighbouring pixels and patterned backgrounds. Nevertheless, on-line performance was still limited to high end processors on personal computers.

In the case of mobile devices, methods often require additional information, such as, using multiple hand-held devices [JS07], or mapping textured background planes [KSY13]. Indeed, *Clayvision* [TP12] was configured to remove original in a set of pre-established locations. *Shadow Retargeting*, uses a background observation method [Mor17], similar to [TP12], in which the user captures the background of the scene before initializing our method.

2.3. Augmented Reality Shadows

An early offline method used observed shadow brightness to recover the illumination of the scene [SSI99] (see Figure 1 b) for such an example). Real-time shadow volumes were prepared manually in [HDH03], to demonstrate the various configurations of virtual and real shadow caster and receiver interactions, but did not address shadow overlap. Jacobs et al [JNA*05], addressed shadow overlap, but did not consider more than one light source or the general light environment. The 3D printed *Shading Probe* captures shadow behavior into a piece-wise constant spherical harmonics basis representation [CMNK13] which is directly applied to rendered diffuse global illumination. However, it remains difficult to design a single 3D printed shading probe which captures both hard and soft shadow behaviour from general illumination environments. Castro et al [CFV12] depict an AR soft shadow estimated using a reflective light probe sphere, but this method only recovers a single light source and shadows cast by static objects. In our light environment estimation, we are similar to Nowrouzezahrai et al [NGM*11] where they also apply a light probe sphere and factorise the primary light direction. Here, we focus on shadow reconstruction through our retargeting method with a variety of estimated real and accurate synthetic light environment scenarios ranging from single point light illumination to area sources and natural illumination environments.

2.4. Stop Motion Animation

Stop motion animation techniques can be a demanding process for movie productions, which can lead to many man years of effort [CKM17]. With additive manufacturing methods, puppets and faces can be produced *en masse* to reduce effort and increase variety [Eme15].

An approach to solve this was introduced through a hand-driven video-based interface [ZW13]. It created automatic temporal interpolation between key-frames in a two-phase based capturing
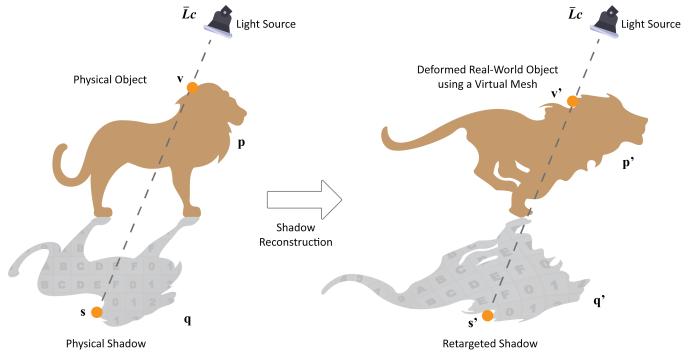

Figure 2: *p defines the real-world static object. q its physical projected shadow. v shows a vertex from the coincident overlaid mesh. s is (v) being projected to the floor. p' the deformed real-world object using a mesh. q' the retargeted projected shadow. v' is the same initial vertex (v) being animated over time. s' is v' being projected to the floor.*

and processing work-flow. It achieved smooth transitions between poses, but did not contemplate reconstructing over occluded areas when not present on the initial frame of reference.

A light-field method was also introduced to reduce costs through a static multi-camera system [ZZK*16] providing an approach for speeding up post-processing effects, enabling changes in the depth-of-field, smooth camera moves and upsampling the video's frame-rate. However, this approach does not address the need to manually create in-between deformation frames and requires a calibrated camera array.

3. Shadow Retargeting

Shadow retargeting leverages the constraints of shadow visibility and appearance with known geometry to efficiently steer source shadow samples for retargeted reconstruction with high quality.

Section 3.1 offers an overview of the method. Section 3.2 details the warping of the physical projected shadow. Section 3.3 illustrates our sampling search for shadow reconstruction. Section 3.4 describes the retargeting model. Section 3.5 contains our procedure for shadows with *umbra* and *penumbra*.

3.1. Method Overview

Our approach synthesizes virtual shadows directly by sampling the image that contains the shadow of the object presented in the real world. As the virtual mesh is animated, we warp the real shadow to retarget its appearance to the virtual one. This process approximates the shadow as being directly linked to the projection along the principal light direction of the caster on the receiver. Our method requires to estimate the direction of the n main light sources in the scene. To name a few, approaches like [WS03], in which multiple directional lights are estimated from a single image, or [WPL18], in which an accurate prediction of ambient light is obtained through a neural network, focused on solving this estimation efficiently for Mixed Reality environments. Since light estimation is not the focus

of our work, we assume that we have information about the direction of the n most important shadow casting light sources of the scene. Our method holds plausibly for small proximity displacements and exploits visibility as a smooth function of the given geometry displacement for natural lighting warped reconstruction.

3.2. Shadow warping

Our goal is to warp the shadow from the image to the displaced version. We rely on prior knowledge of the mesh geometry and register it with its 3D printed physical version.

In a first step, we project the mesh vertices in their world space positions along the light direction on the receiver geometry basing our approach on [Bli88]. We then re-project to image space in order to associate each vertex of the mesh to its position on the shadow.

In a second step, we project the mesh to the ground using the position of those vertices in the virtual pose, i.e animated, and use the previously mapped texture coordinates to interpolate the original shadow across the virtual shadow (see Fig. 2).

3.3. Shadow inpainting

The shadow warping step acquires as much as possible of the whole real shadow visible in the image. In the base case, it does not take into account any further occlusion of the shadow (see section 4.6 for overlapping shadow considerations). We can detect occlusion from the mesh since it is registered in the image. We then search for another reference point to synthesize the shadow appearance.

We synthesize a mask of eligible shadow area valid for sampling, by first rasterizing the projected shadow of the real-world object (Fig. 3b), and then rasterizing the object itself to remove its own occlusion of the shadow (Fig. 3a). This results in an eligible shadow area (Fig. 3c).

Figure 3: *a) Mask for physical object pose. b) Mask for physical shadow pose. c) Mask containing the valid shadow area eligible for appearance sampling.*

Where the shadow warp would sample an invalid texel, we perform a discretized concentric ring search to find the closest valid (i.e unoccluded shadow) texel (Fig. 4). The search samples all the candidates for each iteration at a time. It stops once it finds a candidate that is valid. In the case of having multiple eligible candidates, the one closest to the invalid texel is sampled. To compensate for appearance discontinuities resulting from the inpainting process, we perform a smoothing pass using a box linear filter. Samples outside the shadow are discarded in this pass as we assume a uniform light source.

3.4. Retargeting model

By retargeting shadows, the albedo of the receiver can change over non-regular backgrounds. To take into account materials with non-uniform albedo, we need to relate the outgoing radiance of the receiver in shadow with its radiance in light. The calculated ratio is applied to a new point in the retargeted shadow. For both reference points, in shadow and light, we assume that there is no emissive radiance or subsurface scattering events. We express the outward radiance (L_o) following derivation from the rendering equation 1 [Kaj86].

$$L_o = \int_\Omega f_r(\mathbf{x}, \omega_i, \omega_o) L_i(\mathbf{x}, \omega_i) (\omega_i \cdot \mathbf{n}) \, d\omega_i \qquad (1)$$

$f_r(\mathbf{x}, \omega_i, \omega_o)$ is the BRDF, the proportion of light reflected from ω_i to ω_o at position $\mathbf{x}$. $L_i(\mathbf{x}, \omega_i)$ is the radiance coming towards $\mathbf{x}$ from direction ω_i. $\omega_i \cdot \mathbf{n}$ is the incidence angle factor. ω_i is the negative direction of the incoming light.

Assuming a diffuse BRDF (f_r), the incoming light is scattered uniformly in all outgoing directions. Therefore, in this case, the BRDF does not depend on the incoming (ω_i) and outgoing (ω_o) light directions and becomes a constant determined as ρ (eq. 2).

$$f_r(\omega_i \rightarrow \omega_o) = \rho \qquad (2)$$

Given our diffuse reflectance (ρ_d), we are constrained to a constant subset of full BRDFs. The diffuse BRDF and diffuse reflectance are related by a π factor. (eq. 3).

$$\rho_d = \pi\rho \qquad (3)$$

Since we cannot evaluate the exact irradiance of the real-world scene without accurate light estimation, we assume a constant uniform light source in a position ($\mathbf{x}$) between reference points in shadow (E_{i0}) and light (E_{i1}). (eq. 4).

$$
\begin{aligned}
E_{i0} &= L_{i0}(\mathbf{x}, \omega_i) \\
E_{i1} &= L_{i1}(\mathbf{x}, \omega_i) \\
E_{i0} &= E_{i1}
\end{aligned}
\qquad (4)
$$

Therefore, under these conditions, the ratio of outgoing radiance is conserved for any uniform illumination in the same position ($\mathbf{x}$) in light (L_{o0L}) and shadow (L_{o1S}). This ratio is simply computed as the diffuse BRDF in light (ρ_{d0}) divided by the diffuse BRDF in shadow (ρ_{d1}). (eq. 5).

$$L_{o0L}/L_{o1S} = \rho_{d0}/\rho_{d1} \quad \forall E_i \qquad (5)$$

We use this ratio to compute the new point in shadow (L_{o0S}) drawing from the same reference point in light (L_{o1L}). Both reference points are photometrically calibrated and gamma corrected. (eq. 6).

$$L_{o0S} = \rho_{d0}E_{i0} = \frac{\rho_{d0}}{\rho_{d1}}\rho_{d1}E_{i0} = \frac{L_{o0L}}{L_{o1S}}\rho_{d1}E_{i1} = \frac{L_{o0L}}{L_{o1S}}L_{o1L} \quad (6)$$

3.5. Soft Shadows

Soft shadows are characterized by two main parts, *umbra* and *penumbra*. While the *umbra* represents the area in which all the rays emitted by the light source are occluded, the *penumbra* only represents those rays that are partially blocked.
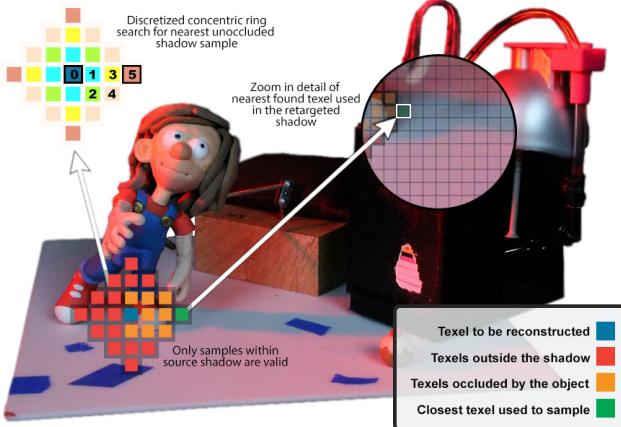
Figure 4: *Discretized concentric ring search algorithm for plausible shadow reconstruction. The red samples are trivially masked and rejected as ineligible. The closest eligible texel, illustrated as a green sample, is used for appearance sampling.*

Algorithm 1: Shadow Retargeting Overview

foreach *light source **in** scene* **do**
    foreach *projected target shadow texel **in** image* **do**
        if *receiver is textured background surface* **then**
            apply retargeting model()
            if *soft shadow edge region* **then**
                return *retargeted PCSS texel*
            else
                return *retargeted shadow texel*
        else if *plain color uniform backdrop* **then**
            if *texel in umbra area (mask 3c)* **then**
                return *warped shadow texel*
            else
                if *soft shadow edge region* **then**
                    return *reconstructed PCSS Texel*
                else
                    return *reconstructed Shadow Texel*

To rasterize a soft shadow, we use a multi-pass approach in which we first reconstruct the *umbra* and then recreate the *penumbra* using Percentage-Closer Soft Shadows (PCSS) [Fer05]. The kernel used to soften the shadow edge is determined by the distance to the first blocker and the solid angle of the light source. To avoid losing texture details in the albedo, we apply a box linear filter to the shadow term L_{o1S} before applying the albedo scaling term $\frac{\rho_{d_0}}{\rho_{d_1}} = \frac{L_{o0L}}{L_{o1L}}$.

4. Algorithm

Our method retargets shadow texels given a corresponding deformation of the geometry for each light source in the scene (algorithm 1). When the shadow is projected onto a textured background, we reconstruct the shadow texel if occluded (section 3.3), and apply the retargeting model (section 3.4). To do so, the outgoing radiance ratio is calculated by comparing the same texel in both light and shadow. Subsequently, this proportion is taken into account for the texel in the retargeted shadow. In the event that the retargeted shadow reveals areas that were initially in the shade, we directly process the previously obtained texel from the observed background image.

In the case of a uniform background on the receiver, we simply wrap the shadow if the texel is inside the *umbra* area (mask 3c). Otherwise, we perform a discretized concentric ring search and reconstruct the shadow (section 3.3). For both cases, if a given texel is inside the *penumbra* area, we perform the reconstruction of the *penumbra* using Percentage-Closer Soft Shadows (PCSS).

4.1. Physical Object Registration

To achieve accurate registration of the physical object, we use marker-less tracking for Augmented Reality [CMPC06] [Vuf17]. These consist of a system based on point clouds that recognizes physical rigid and opaque objects. When detected, this point cloud

registers the origin of the world space coordinate system. The virtual mesh is placed coincident at the origin of this Euclidean space. Using this procedure, we can simulate where the physical object is in our virtual coordinate system. We can estimate the position of the ground plane also from the origin of the Euclidean space, since we assume that the detected object is in contact with the ground.

The digital meshes of the physical objects are created through photogrammetric scans. A stack of digital images from a variety of viewpoints are used in order to obtain the reconstructed mesh and finally those meshes are rigged and animated using Autodesk Maya.

4.2. Shadow Composition Masks

To ease rendering for our Augmented Reality scenario, we use auxiliary masks rendered as textures. These masks, as introduced earlier in section 3.3, are used to detect the position of the physical object in the scene. When creating these auxiliary masks, we will create an instance of the mesh in its initial position and render it to a specific channel bit-mask component. Figure 3a, shows the auxiliary masks in which the position of the physical object is shown. The same procedure is used to segment the physical projected shadow. In this case, we make use of the implementation described in sections 3.4 and 3.5. However, in this case, we return only binary values as shown in figure 3b.

4.3. Real-time Inpainting

In order to compensate for hidden areas from the scene, our approach uses a background observation method [Mor17], similar to [TP12], in which the user previously captures a representation of the background. Masks 3a and 3b are sampled to the shader and evaluated in each fragment over a plane that entirely covers the region in which the physical object would render. Using these masks naively crops out the physical object and its projected shadow from the scene. The segmented area is filled with the previously captured
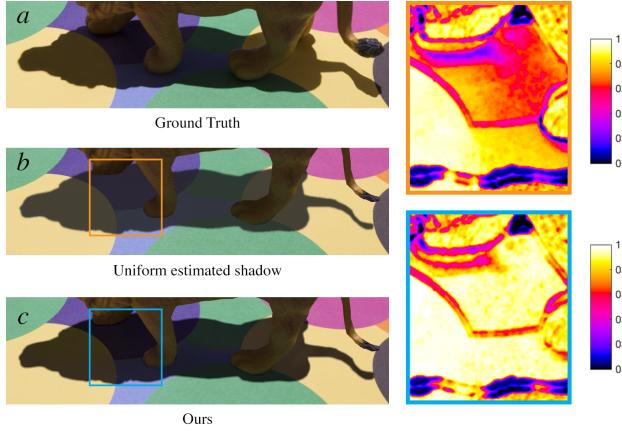
Figure 5: *a) Ground truth shadow generated by a physical occluder. b) Uniform simple estimated shadow. c) Virtual shadow rendered using our method. SSIM comparisons of GT with b) above, c) below showing improved accuracy through shadow retargeting*

representation of the background. Once the image is reconstructed, we perform a linear interpolation between inpainted and unmodified pixels to achieve improved spatial and temporal consistency.

4.4. Object Appearance Retargeting

Following the work of Casas *et al.* [CKM17], we perform object appearance retargeting using texture deformation from a live camera feed video. The initial world coordinates, when converted to screen space, contain the texture information of the physical object. This approach allows for plausible photo-realistic Augmented Reality appearance, since the material, illumination and shading of the object are updated to the virtual mesh in each frame. Any changes applied to the physical object environment will have their expected lighting effects on the virtual mesh.

4.5. Depth Coherent Selection

Our method warps the original shadow of the image to the virtual displaced one. When deforming, several candidates on the warped shadow may be eligible to be rendered in the same projected fragment. If that happens, our method performs depth ordering according to the closest occluder. Therefore, after rasterizing the object, the closest one is represented in the depth buffer. This method preserves the shadow appearance, rendering the closest reconstructed sample when multiple candidates can be represented in the same fragment.

4.6. Handling Multiple Shadow Overlap

When a multiple shadow scenario occurs, each individual shadow has a unique mask in its initial and retargeted position. This allows us to detect shadow regions that overlap in their initial and deformed states. This detection is done at texel level and is evaluated when performing a discretized search for shadow reconstruction. Therefore, when there is an overlapping region, the inpainting

sample is not sampled until all conditions in the search algorithm are fulfilled. For instance, if an overlapping region occurs in the retargeted binary masks, the search algorithm will not provide a sample until it finds a region where those initial masks also overlap. This approach preserves the appearance in cases of multiple overlapping retargeted shadows (figure 6).

5. Results

5.1. Comparisons

Figure 5 shows a comparison between a ground truth shadow generated by a physical occluder, the results using a simple estimated uniform shadow and our method. As can be seen in the Structural Similarity Index Metric (SSIM) comparison, our method is closer to ground truth since it is able to preserve the indirect lighting already present in the scene. Further, our ambient occlusion approximation further reduces error near contact points around the lion's feet. While the 3d object tracking registration and photogrammetry reconstructed model employed has accuracy that results in a visually stable animation in the video frame, slight misalignments result in the lines of higher error where discontinuities edges occur, e.g. in the lower portions of the zoomed in SSIM visualizations of figure 5.

Figure 6 shows a comparison between Ground Truth shadows and our retargeting algorithm. *a-d* demonstrate the real-time capability of our algorithm for Augmented Reality. We perform texture deformation to animate the virtual mesh and shadow samples to reconstruct the occluded areas. Its key frames were 3D printed in order to compare *Shadow Retargeting* with real-world shadows. SSIM results demonstrate good precision using our real shadow retargeting technique, which may be improved with further accuracy of 3D marker-less object registration and tracking. *e-g* show results under High Dynamic Range (HDR) maps [Deb98]. These demonstrate good precision in complex lighting scenarios, where we approximate multiple retargeted soft shadows. *h-j* show synthetic results generated using Unity3D. These show the behaviour of our technique under single and multiple points of light from point to area emitting regions. Our method is able to deliver accurate retargeting under a variety of lighting conditions. SSIM comparisons show high accuracy on *umbra* and *penumbra* regions between Ground Truth and our retargeting approach.

5.2. Performance

Results *(a - d)* were generated with an Apple iPhone X with an output resolution of 2436px by 1125px. We achieved interactive frame rates (25 fps) on this mobile phone. Results *(e-j)* were generated in a 2.8 GHz Quad-core Intel Core i7 with 16 GB of RAM with an image size of 1280px by 720px. We achieved a constant frame rate of 30 fps, which allows us to reconstruct shadows in real time for modern video standards.

As seen in table 1, the primary bottleneck of our system is the discretized search for shadow reconstruction. On average, this takes one third of the render time per frame. However, the number of sample search iterations is typically low, which makes the method suitable for interactive frame-rates on low powered mobile devices.
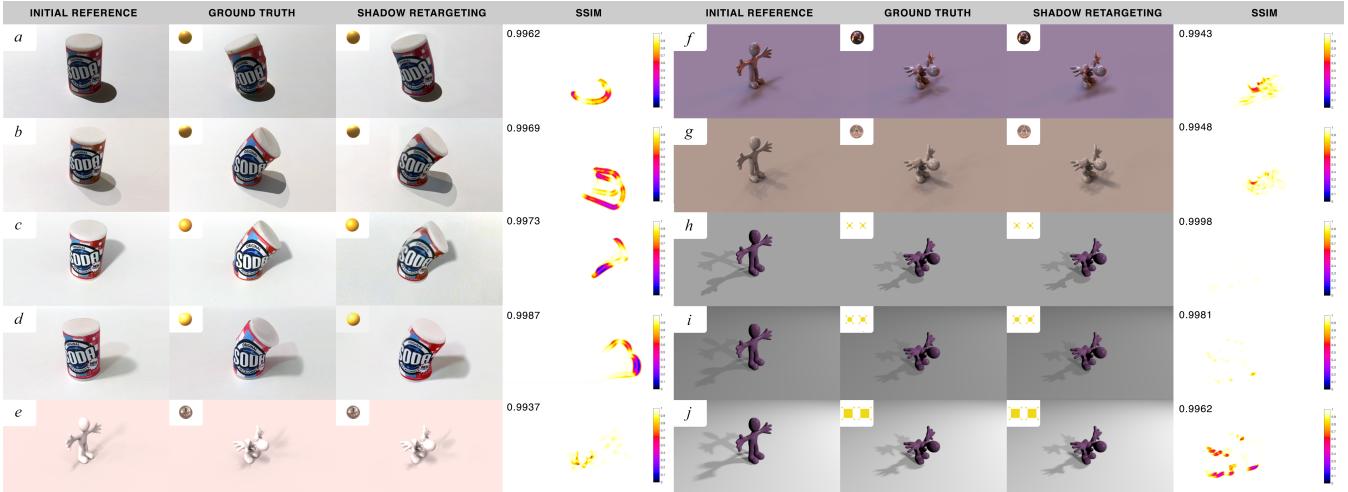
Figure 6: *Comparisons of ground truth with our shadow retargeting scheme. a-d a real 3D printed non-deformed object's shadow retargeted with single/double light sources, and hard/soft shadows. e-g synthetic renders with classic light probes, showing retargeted complex shadows. h-j shows the effect of progression of point to area lighting retargeted from a rendered source image with character in bind pose.*

6. Future Work

Support could be extended for non-planar receivers using multiple adaptive sub-surfaces and depth information. This would introduce some additional complexity to the shadow sample search process and increase the error in the current approach. Reference point selection could be further improved by sampling visibility to obtain a better estimate of the correlation between shadow appearance than simply using the distance to the reference shadow sample. Modelling the shadow appearance on visibility could allow for more advanced techniques to interpolate and even extrapolate the shadow to obtain more consistent results in more complex scenarios. An approach based on bidirectional re-projection could resolve large deformations or topology changes in which the occluder geometry is significantly altered from its physical position [YTS*11]. However the impact of such schemes on the real-time performance of the method should be considered.

The present method relies on the lighting to be dominated by a few principal lights with their casting shadows. We could reconstruct the number of lights, directions, apparent angle and intensity from the shadows. Since there are no models to directly reconstruct those parameters from physical shadows, we would use an iterative search to find the solution of this inverse problem. Additionally, the solution can be refined more efficiently, once the shadows are sufficiently close, by estimating the gradient of the distance function with respect to the solution parameters.

Our method holds plausibly for small proximity displacements. The deformations made under the same visual axis are perceptually more precise. In the case of animated meshes that reveal geometry not present in the reference model, visual artefacts appear due to disocclusions addressed in part by inpainting. Large translations in the virtual object from the physical reference position cause lighting inconsistencies with spatial variations in the environment. Temporary inconsistencies between frames of animation

may appear due to the reconstruction search performed in real-time for occluded areas. These limitations could be surpassed by using a machine learning model that would contain accurate predictions for retargeted poses under large translations or deformations. In addition, our method assumes that the detected object is in contact with the ground. In the case of the object not being in contact with the ground, we could calculate the distance of the object to the floor using an auxiliary marker. Finally, when performing reconstruction in a multiple shadow scenario, we rely on sampling from a physical overlapping region of the shadow to maintain its appearance. Therefore, our method will fail to reconstruct if the physical shadow has no overlapping region that can be used to sample. We anticipate that this limitation can be surpassed by calculating the approximate overlapping appearance using the two projected individual shadows.

Task	Time	Percentage
Object Registration	1,83 ms	4,43 %
Auxiliary masks (Fig 3a, 3b, 3c)	2,59 ms	6,27 %
Object Appearance Retargeting	1,93 ms	4,67 %
Shadow Warping	3,13 ms	7,57 %
Discretized Search	15,78 ms	38,20 %
Uniform Shadow Blurring	7,36 ms	17,82 %
Percentage Closer Soft Shadows	5,27 ms	12,76 %
Background Inpainting	2,43 ms	5,88 %
Scene Rendering	0,98 ms	2,37 %
Total	**41,3 ms**	**100 %**

Table 1: *Time breakdown of a typical frame processed using our Shadow Retargeting method in an Apple iPhone X.*

7. Conclusion

In this paper, we introduced *Shadow Retargeting*. Our work is the first to retarget already present shadows from static objects against their overlaid movement with plausible coherent results. By performing an analytic reconstruction with un-occluded samples, we recreate the deformed appearance of the shadow in real-time. This technique enables plausible photo-realistic in-direct relighting in Augmented Reality. We have shown how *Shadow Retargeting* can be applied to Stop Motion Animations, in which in-between frames of key poses can be computer-generated through Augmented Reality (see the accompanying video). This approach enables a faster and cost-effective way to generate smooth transitions between key frames meanwhile preserving its natural appearance. Additionally, this technique can also be applied to bring animated life to static physical toys. Using texture deformation and *Shadow Retargeting*, static objects may have the power to interact as if they were real.

Acknowledgements

References

[Bli88] BLINN J.: Me and my (fake) shadow. *IEEE Comput. Graph. Appl. 8*, 1 (Jan. 1988), 82–86. 4

[BSCB00] BERTALMIO M., SAPIRO G., CASELLES V., BALLESTER C.: Image inpainting. *Proceedings of the 27th annual conference on Computer graphics and interactive techniques SIGGRAPH 00 2* (2000). 3

[BVSO03] BERTALMIO M., VESE L., SAPIRO G., OSHER S.: Simultaneous Structure and Texture Image Inpainting. *IEEE Transactions on Image Processing* (2003). 3

[CFV12] CASTRO T. K. D., FIGUEIREDO L. H. D., VELHO L.: Realistic shadows for mobile augmented reality. *Proceedings - 2012 14th Symposium on Virtual and Augmented Reality, SVR 2012* (2012). 3

[CKM17] CASAS L., KOSEK M., MITCHELL K.: Props Alive : A Framework for Augmented Reality Stop Motion Animation. In *2017 IEEE 10th Workshop on Software Engineering and Architectures for Realtime Interactive Systems* (Los Angeles, California, USA, 2017). 3, 6

[CMNK13] CALIAN D., MITCHELL K., NOWROUZEZAHRAI D., KAUTZ J.: The Shading Probe: Fast Appearance Acquisition for Mobile AR. *SIGGRAPH Asia 2013 Technical Briefs* (2013). 3

[CMPC06] COMPORT A. I., MARCHAND E., PRESSIGOUT M., CHAUMETTE F.: Real-time markerless tracking for augmented reality: The virtual visual servoing framework. *IEEE Transactions on Visualization and Computer Graphics 12*, 4 (2006), 615–628. 2, 5

[Deb98] DEBEVEC P.: Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *Proc. ACM SIGGRAPH '98 Proceeding* (1998), ACM. 6

[Eme15] EMERSON S.: Visual Effects at LAIKA, A Crossroads of Art and Technology. *ACM SIGGRAPH 2015 Talks* (2015), 2799649. 3

[Fer05] FERNANDO R.: Percentage-closer soft shadows. In *ACM SIGGRAPH 2005 Sketches on - SIGGRAPH '05* (2005), p. 35. 5

[GZMT10] GOFERMAN S., ZELNIK-MANOR L., TAL A.: Context-aware saliency detection. *IEEE Conference on Computer Vision and Pattern Recognition* (2010). 2

[HB12] HERLING J., BROLL W.: PixMix: A real-time approach to high-quality Diminished Reality. In *11th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (2012). 3

[HDH03] HALLER M., DRAB S., HARTMANN W.: A real-time shadow approach for an augmented reality application using shadow volumes. *Proceedings of the ACM symposium on Virtual reality software and technology - VRST '03* (2003). 3

[JNA*05] JACOBS K., NAHMIAS J.-D., ANGUS C., RECHE A., LOSCOS C., STEED A.: Automatic generation of consistent shadows for augmented reality. *ACM International Conference Proceeding Series; Vol. 112* (2005), 113. 3

[JS07] JARUSIRISAWAD S., SAITO H.: Diminished reality via multiple hand-held cameras. *2007 1st ACM/IEEE International Conference on Distributed Smart Cameras, ICDSC* (2007), 251–258. 3

[Kaj86] KAJIYA J. T.: The rendering equation. *SIGGRAPH Comput. Graph. 20* (Aug. 1986), 143–150. 4

[KS00] KANG S. B., SHUM H.-Y.: A review of image-based rendering techniques. Institute of Electrical and Electronics Engineers, Inc. 2

[KSY13] KAWAI N., SATO T., YOKOYA N.: Diminished reality considering background structures. In *2013 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2013* (2013). 3

[LLT*11] LEÃO C. W. M., LIMA J. P., TEICHRIEB V., ALBUQUERQUE E. S., KEINER J.: Altered reality: Augmenting and diminishing reality in real time. In *Proceedings - IEEE Virtual Reality* (2011). 2

[Man01] MANN S.: Fundamental issues in mediated reality, WearComp, and camera-based augmented reality. *Fundamentals of Wearable Computers and Augmented Reality, Lawrence Erlbaum Associates* (2001). 2

[MNZ*15] MAGNENAT S., NGO D. T., ZÜND F., RYFFEL M., NORIS G., ROTHLIN G., MARRA A., NITTI M., FUA P., GROSS M., SUMNER R. W.: Live Texturing of Augmented Reality Characters from Colored Drawings. *IEEE Transactions on Visualization and Computer Graphics 21* (2015), 1201–1210. 2

[Mor17] MORI S.: Diminished Hand : A Diminished Reality-Based Work Area Visualization. *2017 IEEE Virtual Reality* (2017). 3, 5

[NGM*11] NOWROUZEZAHRAI D., GEIGER S., MITCHELL K., SUMNER R., JAROSZ W., GROSS M.: Light Factorization for Mixed-Frequency Shadows in Augmented Reality. In *10th IEEE International Symposium on Mixed and Augmented Reality, ISMAR* (2011). 2, 3

[SSI99] SATO I., SATO Y., IKEUCHI K.: Illumination distribution from shadows. *Computer Vision and Pattern Recognition (CVPR)* (1999). 3

[STR*05] SETLUR V., TAKAGI S., RASKAR R., GLEICHER M., GOOCH B.: Automatic image retargeting. *ACM International Conference on Mobile and Ubiquitous Multimedia* (2005). 2

[TP12] TAKEUCHI Y., PERLIN K.: ClayVision: the (elastic) image of the city. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2012), 2411–2420. 2, 3, 5

[VTP*10] VAQUERO D., TURK M., PULLI K., TICO M., GELFAND N.: A survey of image retargeting techniques. *Proc. Applications of Digital Image Processing XXXIII* (2010). 1, 2

[Vuf17] VUFORIA: Vuforia Object Recognition, 2017. 2, 5

[WPL18] WEBER H., PREVOST D., LALONDE J.-F.: Learning to estimate indoor lighting from 3d objects. *CVPR* (2018). 3

[WS03] WANG Y., SAMARAS D.: Estimation of multiple directional light sources for synthesis of augmented reality images. 3

[YTS*11] YANG L., TSE Y.-C., SANDER P. V., LAWRENCE J., NEHAB D., HOPPE H., WILKINS C. L.: Image-based bidirectional scene reprojection. *ACM Trans. Graph. 30*, 6 (Dec. 2011), 1–10. 7

[ZW13] ZHENG X., WANG L.: A Video-based Interface for Hand-Driven Stop Motion Animation Production. *IEEE Computer Graphics and Applications* (2013). 3

[ZZK*16] ZILLY F., ZIEGLER M., KEINERT J., SCHOBERL M., FOESSEL S.: Computational Imaging for Stop-Motion Animated Video Productions. *SMPTE Motion Imaging Journal 125*, 1 (2016), 42–47. 3

[ZZX15] ZHANG L., ZHANG Q., XIAO C.: Shadow remover: Image shadow removal based on illumination recovering optimization. *IEEE Transactions on Image Processing 24*, 11 (Nov 2015), 4623–4636. 2