Benchmarking surface selective vacuum ultraviolet and thermal postprocessing of thermoplastics for ultrasmooth 3-D-printed micro-optics

Robert Kirchner,a,b Nachiappan Chidambaram,b and Helmut Schiftb,*

aTechnische Universität Dresden, 3D Mesoscopic Systems, Dresden, Germany
bPaul Scherrer Institute, Laboratory for Micro and Nanotechnology, Villigen PSI, Switzerland

Abstract. State-of-the-art, polymeric, refractive micro-optics simultaneously require an ultrasmooth three-dimensional (3-D) surface and a precise geometry for excellent optical performance with minimal stray light. In earlier work, we have established a surface finishing process for thermoplastic polymer master structures that is only effective on the surface and does not affect the designed optical geometry, thus enabling polishing without touching. Therewith, the high curvature corners of a 50-μm-tall optical diffuser device were maintained while the surface roughness was reduced to about 10-nm root mean square. For this, 3-D master structures were first fabricated by direct write laser-lithography with two-photon polymerization. The master structures were replicated into poly(methyl methacrylate) through a poly(dimethyl siloxane) intermediate replication stamp. Finally, all structures were surface-polished by selective high-energy photon exposure and thermal postprocessing. In this work, we focus on the comparison of the surface smoothening using either postprocessing or dedicated direct writing strategies. For this comparison, strategies for modifying the exposed voxel size and the writing discretization being the primary source of roughness were tested by sweeping the laser exposure dose for two different resist materials and objectives. In conclusion, the postprocessing smoothening resulted in a lower roughness compared to a direct writing strategy—even when 50-nm vertical discretization steps were used—and still enabled 10 times shorter writing times. © 2018 Society of Photo-Optical Instrumentation Engineers (SPIE)

Keywords: microlens array; glass transition temperature; surface roughness; polymerization; direct laser writing; two-photon; multiphoton.

Paper 171144SSP received Jul. 21, 2017; accepted for publication Nov. 13, 2017; published online Jan. 6, 2018.

1 Introduction

Advanced microlithographic methods offer completely new and before practically impossible three-dimensional (3-D) optical designs to be realized. In addition, these methods enable a quick turnaround by fast-prototyping designs for later high volume manufacturing (HVM). New, compact optical designs with refractive and diffractive properties are required in advanced, integrated optical applications such as cameras, optical sensors, and diffusers, or movement tracking surveillance sensors. Since these components are highly integrated and have a small footprint, they have to be fabricated in 3-D lithographic methods enabling high structural fidelity, precision, and resolution.

The fabrication schemes include 3-D methods such as grayscale lithography, which tunes the resist profile by the exposure dose, and especially, two-photon polymerization (2PP) lithography, which builds 3-D structures out of thin discrete layers. Grayscale approaches are either based on masks with varying transmission grades (grayscale masks), direct writing methods, or multiple exposures. While the first one is well established for i-line optical lithography (365 nm wavelength), higher resolution, direct writing methods are based on sequential scanning of photons, electrons, or ion beams on predefined trajectories through the material. Traditional direct laser writing (DLW) lithography using single-photon absorption with a lateral pitch resolution of about 700 nm is well suited for refractive micro-optics. However, its constraint to resist films with moderate thickness (typically around 30 μm) limits its use for freeforms and complex designs beyond this thickness. This is mainly due to the fact that the entire film is exposed and absorbs photons along the beam path when using single photon absorption processes. Similarly, the higher resolution grayscale methods based on electron and ion beams are mostly suitable for diffractive optics with film resist thickness below 10 μm for electrons and even thinner for ions. A true 3-D method such as 2PP-DLW allows for full freeform designs without restrictions on the film thickness and the size of the micro-optics. Exemplary designs realized by 2PP are micro-prisms, single-tier microlenses, multitier lenses, and optical freeforms.

The typical limitation of 2PP-DLW is the writing time and thus writing area and build-height. This is because of the highly confined region in which the nonlinear, multiphoton absorption takes place, which enables very high-resolution 3-D writing but also requires extensive layering for bulky structures. The laser beam focus (a spheroid called voxel) is scanned through a viscous material and polymerizes the material along its trajectory, which means that the structure is built from overlapping trajectories that need to be laterally adjacent and stacked on top of each other.
For HVM, micro-optical elements are typically replicated in large quantities by polymer molding techniques such as hot embossing, injection molding, casting, or UV-assisted methods. Typical industrially relevant thermoplastic polymers are commodity plastics such as poly(carbonate), poly(methyl methacrylate) (PMMA), or other thermoplastic polymers with appropriate thermomechanical, optical, and chemical properties. Replication methods can achieve faithful replication of sub-10-nm surface topographies. A fabrication method for a structure intended as a mold for replication would thus need the mold to be as smooth as possible. Very often, the pattern methods used to create the originals entail a trade-off between processing time and the resolution as well as the roughness to be achieved. With current 3-D lithographic methods, direct fabrication of the required ultrasmooth surfaces can lead to substantial increase of writing times, which are prohibitive for many large area applications. For cost efficient HVM, even very fast direct writing processes being available for large area originals would probably avoid step-and-repeat enlargement of small-scale original patterns but would still require replication.

This work deals with methods that enable the fabrication of microlenses with ultrasmooth, optical grade surfaces. Instead of improving surface roughness by time-consuming and design-specific direct-writing strategies of 3-D lithographic methods, we chose a postprocessing method based on surface selective thermal reflow. We quickly recapitulate this method for the modification of polymer surfaces being published earlier and especially compare in this work the two strategies of postprocessing and direct-writing for surface smoothening. The principle of smoothening is based on a process known as thermally activated selective topography equilibration (TASTE). TASTE uses high-energy radiation such as electrons or photons, such as those employed in this work, to selectively modify the glass transition temperature of a defined region only. Upon a global thermal treatment, the modified region will soften and undergo surface-tension driven material relocation while the unmodified regions remain unaffected. This enables a surface smoothening and thus is a contactless polishing.

2 Roughness in Micro-Optics

2.1 Importance of Roughness Mitigation

Surface roughness influences the performance of optics by reducing image contrast or signal-to-noise ratio, and a maximum allowable roughness limit must be specified. Here the signal-to-noise ratio is defined by specular reflectance over scattering. The roughness specification includes the maximum allowed roughness value for a specific optical wavelength range and the spatial bandwidth over which it is to be met. While for micro-optics this is often not specified, the surface roughness is more critical for retroreflectors and mirrors. In general, good imaging optics require an overall surface roughness about 1/100th of the wavelength used to achieve low stray light, high resolution, and good imaging contrast. For relevant visible wavelengths of 400 to 800 nm, this would be a roughness of 0.7·σq, the surface could be as rough as 35 nm. Considering the 1/100th rule mentioned above, this is a very loose tolerance for optical surfaces, and normally this cannot be accepted. Empirical expression relating reflection efficiency, originally of a conductive mirror, with roughness and wavelength was derived in Ref. to be

$$\frac{R_q}{R_0} = e^{-(\frac{\sigma_q}{\lambda})^2},$$

(1)

or for inclined incidence with angle $\alpha$

$$\frac{R_q}{R_0} = e^{-(\frac{\sigma_q \cos \alpha}{\lambda})^2},$$

(2)

where $R_q$ is the specular reflectance of the rough surface and $R_0$ is the specular reflectance of an ideal surface with no roughness. This relation also imposes very strong restrictions on the roughness. Note that RMS roughness alone does not give a full indication about the surface finish it is also the distribution of spatial frequencies of the features.

2.2 Roughness Reduction: Writing Strategies Versus Postprocessing

Based on the principle of 2PP scanning a fine laser focus with 3-D trajectories through the material to be exposed, there are three major concepts to minimize the surface roughness of 2PP written structures (Fig.):

- process-based methods such as increasing the voxel size by varying the objective, the laser power (LP), or the resist behavior, thus enabling to write trajectories with larger polymerization extent and smoother transitions;
- design-based writing strategies such as increasing voxel overlap, adaptive vertical discretization based on the curvature of the feature to be written, writing contours either with curved out-of-plane, or by spiraling trajectories for the outer shell, thus enabling to write trajectories with larger and design-optimized overlap;
- postprocessing methods such as the one described in this work that largely decouples the process- and design-based strategies from the smoothening process.

The roughness, for a typical structure made by 2PP-DLW, would be mostly constituted by the height of the mounds formed by adjacent voxels, which is measured against the
dip between them. Due to the threshold of 2PP, the mounds are constituted by unblurred voxels. For different slopes and discretizations, this regular undulation can be deduced from both voxel size and overlap. Additional roughness comes from stitching errors and material inhomogeneities but are not subject of this research. In the following, the first two strategies will be discussed in detail.

In principle, the voxel size can be either decreased or increased to reduce the roughness. While a smaller voxel size means writing with a much finer probe and requires a much finer discretization; a larger voxel size is far more practical. Using the same level of discretization, a larger voxel size means an increased voxel overlap and thus much smaller layer-to-layer steps and thus less surface roughness [Fig. 2(a)]. In contrast, smaller voxel size means less overlap and more roughness asking for improved discretization. The voxel size can be tuned by the objective, the laser power or the resist chemistry. By far the largest effect on the voxel size stems from the used focusing objective. Low numerical aperture (NA) and small magnification objectives enable a large voxel size as required for smooth surfaces and a quick volume filling. Switching between objectives with different NA allows for adjustment between volume writing (low NA) and high fidelity patterning (high NA) assuming an accurate enough overlay is possible. Additional voxel enlargement can be achieved based on the polymerization threshold of the resist. With increased laser power, the volume above threshold is increased leading to large voxels as shown in this work. For constant laser power, lowering the polymerization threshold by synthesizing specific resists will have the same effect. Furthermore, using resist with an increased diffusion range of the active radicals required for polymerization will also lead to larger voxel sizes.

Increasing the voxel overlap by finer discretization can reduce the roughness to some extent [Fig. 2(b)], but resists tend toward microexplosions due to the dielectric breakdown of the polymerized material by the accumulated dose and absorbed energy in the overlap region. In any case, a finer discretization equals a longer processing time. With a variable vertical discretization, needed for surfaces with varying slope angles, overlapping can be enhanced in areas of high discretization induced roughness, but one needs to dynamically tune the dose to avoid microexplosions. Out-of-plane trajectories can give some improvements but this is essentially shifting the roughness problem from “z” axis to “x” or “y” axis for circular segment scanning along the outer structures surface as it is useful for microlenses for example. This contouring can also be used to only write the outer shell of convex structures in a spiral path to give a smooth and continuous trajectory. In general, such methods, which avoid multiple start- and end-points of trajectories on the outer surface, minimize roughness, too. This is because both start- and end-points of the laser trajectory are never ideal in terms of shape and laser parameters due to machine-based restrictions. While piezo drives of current 2PP-DLW systems allow for the scanning of trajectories in all directions, fast writing is often achieved when the in-plane scanning in lateral direction is done via fast galvanic scanners. This allows fast layer-by-layer writing for which the layering in the vertical direction becomes the time-limiting factor of the final 3-D design and continuous, full 3-D trajectories need to be written either with piezoscanners or synchronized piezo- and galvano-meter-scanners. Because of the fast writing, the galvanometric-scanning is often the method of choice, which limits the strategies described before to some extent. Switching between writing modes is often limited, too. One particular strategy is the core-shell method: only the shell of a structure is written with high precision while the inner part of the contoured regions can be written in a different way or even left unexposed. Then, the final structure can be just flood exposed for volume polymerization by single photon absorption to reduce the total process time. This post-2PP flood exposure is only possible if the shell is dense and does not allow the liquid polymer precursor to leak out during wet development.

**Fig. 1** Principle of 2PP with a tightly focused volume element (voxel) where the nonlinear multiphoton process occurs that leads to polymerization of UV-sensitive resist with a certain polymerization threshold as used in this work. Depending on the exposure optics, the voxel is a more or less extended ellipsoid.

**Fig. 2** Comparison of three main alternatives for surface smoothening being either (a) variation of voxel size (with exposure optics/exposure dose/resist chemistry) or (b) by writing strategies ranging from voxel overlap adjustment to various design-specific solutions or (c) by postprocessing. Apart from the resist with diffusive components, the trajectories are constituted by the assembly of unblurred voxels.
It is noteworthy that differences in important material properties such as refractive index (density) and thermal expansion coefficient were reported in dependence of the laser exposure intensity. High-resolution features and high aspect ratio structures will suffer the most from writing strategy approaches that will remain reserved for special geometries where a postprocessing is not possible due to hidden but functional structures such as in multitier lenses. We believe that a universal postprocessing is the appropriate strategy for attaining a super smooth surface while having limited change in the surface profile alongside acceptable writing times. It works particularly well for structures that can be afterward replicated, i.e., structures with no undercuts and masked features. Then also exposure schemes can be applied in which the surface is accessible by radiation or other directed probes. In contrast to some quite specific solutions, postprocessing as shown in this work can be applied to a broad field of structures and applications just using a surface "polishing" process after the writing. Since the process needs to be based on noncontact and should not involve a coating step, it requires a surface selective modification of the polymer in a depth that is comparable to the surface roughness to be smoothed out. Such surface selective modification is possible by exposure of the polymer with high-energy radiation. Similarly to the process established for grayscale electron beam lithography structures, the TASTE process, we established a reduction of the molecular weight of the polymer and thus the glass transition temperature with a high degree of localized vertical confinement. Since the original structures are used mostly as templates for molding, only the surface of the lenses needs to provide the roughness needed for optical performance. The bulk, however, will determine other structural issues such as the overall fidelity and bows (crown bow) due to shrinkage and a possible collapse of structures.

3 Materials and Methods

3.1 Process Flow for Ultrasmooth Surfaces

The TASTE process is best explored for linear thermoplastic polymers such as PMMA and did so far not show satisfying results with highly cross-linked materials such as the typical 2PP-DLW-negative tone resists. It was not possible to achieve a similar confined layer with a pronounced reflow behavior in such a resist. Due to the highly cross-linked dimeromer behavior, the 2PP-DLW resist rather was sublimed upon high-energy radiation than obtaining a thermoplastic surface film. Therefore, a two-step replication process that allowed transferring the lenses from the cross-linked proprietary material optimized for 2PP-DLW to PMMA (Fig. 3) was established. The lens structures were written by 2PP-DLW (Fig. 3(b)). Subsequently, the patterns were transferred into UV-curing poly(dimethylsiloxane) (PDMS) by casting and UV curing (Fig. 3(c)). Finally, the lens patterns were transferred into a thick PMMA film using thermal nanoimprint lithography (Figs. 3(d)–3(f)). The smoothening technique detailed later was applied to the PMMA. For further upscaling, the process chain involves further copying of the smooth lens structures by molding techniques into the desired product material.

3.2 Mastering by Two-Photon Lithography

2PP-DLW was performed using a Photonic Professional GT (Nanoscribe GmbH) in IP-Dip and IP-S (both commercial materials from Nanoscribe GmbH) with moderate scan speeds of 20 mm/s using the galvanometric scanners and low pulse intensities around 1 nJ. For determining the voxel size in dependence of the exposure dose, the laser intensity was swept in a broad range from minimal intensity required for polymerization to the maximum laser intensity (machine based or defined by onset of microexplosions). The used laser is a 780-nm fs-fiber laser with an average power above 120 mW. The parameter sweep was done for both the high-resolution resist IP-Dip as well as the medium-resolution resist IP-S. Both resists were tested with a 63× NA 1.4 high-resolution objective and a 25× NA 0.8 low-resolution objective. For assessing the maximum voxel width, all voxels were directly placed on the substrate surface using the precision autofocus to avoid partially burying the voxel in the substrate. To make sure the substrate surface is found correctly, the adhesion promoter OrmoPrime (n = 1.34 to 1.35, microresist technology GmbH) was used to give a sufficient interface signal for the IP-S resist (n = 1.48).

![Fig. 3 Schematic of the process flow starting from (a) definition of master structures in the commercial resist IP-Dip by 2PP-DLW, (b)–(d) replicating the master (stamp) into a UV curing material to generate the intermediate stamp with negative tone and (e)–(f) replicating the intermediate mold structure into a thermoplastic material by thermal NIL. Modified and reproduced by permission. © 2016 AVS.](https://www.spiedigitallibrary.org/journals/Optical-Engineering)
on the fused silica ($n_{780\text{nm}} \approx 1.45$) substrates. Fused silica was used as substrate to avoid significant back reflections as would be present on silicon due to the large refractive index mismatch leading to an influence on the voxel size. For IP-Dip ($n_{780\text{nm}} \approx 1.52$), no special substrate preparations were required. Voxel height was measured using a stylus profilometry (Dektak 8, Veeco Instruments Inc.).

For fast writing of the large microlens array (MLA), 500 nm slicing distance between adjacent vertical layers was used and the lateral voxel spacing was about 200 nm. This resulted in an in-plane roughness of about 20 nm RMS and respective steps due to the coarse discretization. For comparison, small $3 \times 3$ MLAs were exposed in IP-Dip and IP-S with different slicing distances ranging from 500 to 50 nm. In all cases, a bulk filling method was used. All MLA structures were written on fused silica substrate treated with an adhesion promoter (APS1 or OrmoPrime, microresist technology GmbH) for improved adhesions during the replication process. After exposure, all structures were developed in propylene glycol methyl ether acetate for 10 to 30 min and in isopropyl alcohol for 10 to 30 min and let to be dried under room atmosphere.

### 3.3 Replication by Thermal Nanoimprint Lithography

For replicating the first copy as negative from the IP-Dip master, a UV-curing PDMS (X-34–4184, Shin-Etsu Chemical Company, Ltd.) was drop-casted on the original IP-Dip structure, backed by a Borofloat® glass and UV cured. After overnight storage, the UV-PDMS-glass stack was demolded. The UV-PDMS is flexible enough that high fidelity structures are prevented from being broken but is at the same time hard enough that all structures can be replicated without losses. For copying into the thermoplastic film as positive of IP-Dip, PMMA120k (microresist technology GmbH) was spin-coated in four subsequent steps giving each about 10 μm thickness to obtain the required 40-μm film thickness on a Borofloat glass substrate. Both, the IP-Dip original on fused silica as well as the UV-PDMS replica on Borofloat were treated with an antisticking coating in a chemical vapor deposition process at room temperature following a low power RIE O$_2$ plasma activation with 20 sccm O$_2$, gas flow, 20 mTorr chamber pressure, 20 s plasma time, and 20 W RF power (Plasmalab 80, Oxford Instruments). The replication into PMMA was performed in a Jenoptik HEX 03 press at 2.5 MPa and 180°C with a holding time at high pressure of 10 min. The process window for this robust PDMS-based replication is rather large in terms of imprint pressure and temperature and preserves the original roughness with only minor changes which makes the replication process very appealing as an industrial process.

### 3.4 Surface-Confined Modification by 172-nm VUV Exposure

The PMMA patterns were exposed for 30 s with a vacuum UV (VUV) intensity of about 16.3 mW/cm$^2$ at 172 nm wavelength in a flood exposure tool (EX-mini, Hamamatsu Corp.) to modify the top surface of the structures only (Fig. 4).

Regarding the effect of VUV photons, ion bombardment, and heating on the roughness of resist materials, authors of Ref. found that only a minimal effect is visible due to sole photons or ions exposure. Only the simultaneous photon and argon ion impact at elevated temperatures resulted in a surface roughening. As the used VUV exposure was operated under ambient conditions in air atmosphere, no significant roughening by the VUV exposure itself was expected and found. There are alternative methods (Table 1) for surface modification. They do not only differ in penetration depth but also in their way of interaction with the polymer surface. In Ref. Choi et al. investigated the efficiency of different radiation probes for PMMA molecular weight reduction. They found that x-ray (0.8 to 1.8 keV) was more efficient than electron beam (25 keV) and deep UV (4 to 6 eV) in terms of the material effects leading to main chain scission. They explained this by different reaction pathways. A highly efficient exposure method creates a relatively large amount of main-chain scission events per unit length of a PMMA molecule in comparison to the amount of ester group removal and C==C double bond formation.

One important aspect is the amount of material that is removed during the modification process due to the generation of volatile compounds directly during exposure or during heating. We measured an ablation of about 25 to 50 nm for VUV exposure, with a 200- to 500-nm-thick skin of modified material remaining on the surface. This is tolerable for microstructures and this seems to be the minimum achievable with VUV and EUV exposure. Other probes are much more aggressive. The benefit of VUV is the fact that it can be operated in ambient atmosphere and that...
large area lamps enabling homogeneous irradiation are available. Further materials details, especially on penetration depth, for VUV and EUV exposure can be found in Ref. 28.

3.5 Selective Thermal Reflow by Global Heating

After exposure, the lenses were annealed in a top-heater setup (Fig. 3) at a heater temperature of 140°C for 15 min. Although TASTE is based on a material contrast that enables selective reflow with global thermal treatment (e.g., in an oven at a constant temperature), this setup allows for generating a temperature gradient from the top of the lenses toward the base, which additionally supports the concept of surface confined softening. After reflow, the samples were immediately cooled to room temperature, coated with 5 to 10 nm chromium and inspected by scanning electron microscopy (SEM, Zeiss Supra 55 VP, Zeiss Microscopy). The created contrast in $T_g$ allows a selective softening of the top film versus the bulk material. During the softening, the remaining small polymer chains in the top film become much more mobile compared to the bulk chains. The surface tension of the softened polymer film drives a roughness equilibration toward smoother surfaces, and thus the material of voxel-based mounds will dislocate sideways and preferentially fill the dips between the mounds.

4 Results and Comparison of Methods

4.1 Mastering by Two-Photon Lithography: Variation of the Voxel Size and Slicing Distance

The voxel size was found to be tunable in a relatively large range independent of the resist materials tested just by adjusting the laser intensity (Fig. 4). Due to the constant writing speed, the exposure dose depends only on the laser intensity or power. As expected from the literature, the focusing objective with the higher NA resulted in a smaller voxel size compared to the low NA objective. The voxel width of the IP-S resist could be tuned between 170 and 395 nm as well as between 345 and 668 nm for the NA 1.4 and the NA 0.8 objective, respectively. For the same two objectives, the voxel width for the IP-Dip resist was found to be between 168 and 362 nm as well as 245 and 486 nm. This is in agreement with the purpose of the IP-Dip resist for high-resolution applications and the IP-S for medium resolution and large volume fabrication. In both cases, the voxel size scaled linearly with the laser power increase. However, it is interesting to note that the voxel increase for the IP-S resist is more significant for the 25x NA 0.8 objective. The voxel height for IP-S was related to the width by a factor of 3.3 ± 0.6 for NA 1.4 and 9.5 ± 1.2 for NA 0.8. For IP-Dip, it was 4.0 ± 0.6 for NA 1.4 and 11.9 ± 1.0 for NA 0.8. This is also in good agreement with the literature values of about 8.0 (NA 0.6) and 4.2 (NA 1.4) as well as 3.1 to 3.4 (NA 1.4).

![Fig. 5 Typical process latitude $\Delta T_1 - \Delta T_2$ for the thermal reflow process: (a) balancing glass transition temperature of the bulk $T_g,\text{bulk}$ of 120°C and the process temperature $T_{\text{reflow}}$. For a reflow of a surface skin layer down to 100 nm depth only, the exemplary glass transition temperature $T_g,100\text{nm}$ at this depth has to be less than 40 K below $T_g,\text{bulk}$ and thus at maximum 80°C. (b) The top-heater setup takes into account small differences in temperature between the top and the bottom of the polymer structures and enables better control of the process temperature.](https://www.spiedigitallibrary.org/journals/Optical-Engineering)
considering that the authors used another resist material and a different setup, which influences these values.

For the laser power of about 65% used for MLA manufacturing in this paper, the voxel size with a 63×NA 1.4 objective for the IP-Dip is typically about 260×260×1180 nm³, whereas for the IP-S it is enlarged to 290×290×1100 nm³. For the 25×NA 0.8 objective, a voxel size was about 300×300×3380 nm³ for the IP-Dip and about 380×380×3960 nm³ for IP-S, including the blurring due to the chemical nature of the latter resist. These measurements enable the estimation of the following voxel overlap (Table 2). The results are very similar for both resists and mainly depend on the respective objective. In the given setup, which was not further optimized for writing speed, the writing time for a 3×3 MLA was about 5 min for 500 nm slicing and 50 min for 50 nm slicing. Such MLAs constituted from many identical lens elements are used as refractive diffusers for which the field of view and the uniformity is directly linked to the surface roughness and structural fidelity.

Figure 8 displays SEM micrographs of a 3×3 MLA manufactured by 2PP-DLW displaying the effect of variation of voxel size [Figs. 8(a)–8(c)] and discretization distance defined by vertical slicing [Figs. 8(f)–8(c)] on the surface roughness. All lenses in Figs. 8(a)–8(c) were written in the same time (34 s per single lens or about 5 min for the 3×3 MLA), i.e., with the same scanning speed. The comparison shows that an increased voxel overlap due to the medium resolution objective reduces the roughness significantly [Figs. 8(a) versus 8(c)]. Using an additional overlap increase, due to a mainly lateral voxel increase by a more diffusive resist curing behavior, reduces the roughness even more. The blurring due to the changed resist is clearly visible in the central part of the lens [Figs. 8(b) versus 8(c)].

Further improvements in the surface roughness can be achieved by reduction of the vertical discretization (slicing) distance. For a reduction from 500 to 300 nm, there were still details of layering remaining in the deep sags of the concave lenses [Figs. 8(c) versus 8(d)]. Finally, for 100 and 50 nm slicing almost no roughness from layering could be seen in the SEM micrograph but some, important to note, material intrinsic and granular surface roughness remains even for 50 nm slicing [Fig. 8(e)]. However, the writing time of the structure in IP-S with 50 nm slicing is 10 times (one order of magnitude) larger than that of the original structure with 500 nm slicing in IP-Dip. Such differences in writing times will result in huge processing times, e.g., if lens arrays with a few millimeters of diameter need to be exposed.

4.2 Replication by Thermal Nanoimprint Lithography

For testing the surface smoothening in PMMA, first the IP-Dip master patterns with 500 nm slicing was faithfully replicated into PMMA including the step-to-step roughness due to the coarse slicing (Fig. 8). The PMMA replication resulted in a relatively thick residual layer, which supported together with the flexible PDMS mold a defect free release of the microstructures, the high-aspect ratio features as well as the nanotexture on the microlens surface. Some material shrinkage is visible in the MLA sidewalls [Figs. 8(a) versus 8(c)].
4.3 Surface-Confined Modification for Selective Thermal Reflow During Global Heating

VUV radiation of 172-nm wavelength is strongly absorbed in PMMA with a linear absorption coefficient of about 5 \( \mu \text{m}^{-1} \). Thus, the relative VUV intensity drops to \( 1/e \) already about 200 nm below the PMMA surface and most of the chain-scission and molecular weight reduction will occur there. The molecular weight \( M_W \) significantly influences the glass transition temperature \( T_g \) once it drops below a certain critical \( M_W \) being about 11 to 15 kg/mol for narrow \( M_W \) distributed PMMA. Comparison of wet development ratio of electron-beam to VUV-exposed PMMA indicates that the \( M_W \) in the top PMMA region is well below 10 kg/mol. The lower \( M_W \), the lower is the \( T_g \) and thus the lower is the viscosity and the higher is the ability to flow at a given process temperature above \( T_g \). First experiments based on the change of the coefficient of thermal expansion and white light reflectometric monitoring of the film thickness during a temperature ramp indicate that the \( T_g \) is below 50°C for the top of the VUV exposed PMMA. The \( T_g \) of the unexposed PMMA is about 120°C with a \( M_W \) of about 100 kg/mol. This means the mobility of the polymer chains on the top surface during the 140°C reflow process is significantly higher than that of the bulk material. Due to this
mobility gradient, the top surface will dislocate faster than the underlying layer of nonmodified PMMA. In addition, the top heating setup further reduces the temperature at the base of the MLA, all together avoiding a feature collapse. Since this temperature gradient will depend on the PMMA thickness and the heating time, it can only be roughly estimated to be between 2°C and 10°C. It has to be mentioned that the confined material displacement or reflow is also dependent on the structure size. A much smaller structure, such as a grating would undergo a significant reflow at the used conditions. However, the micrometer-sized bulk geometry and even tips between the adjacent microlenses were not visibly altered. Therefore, the reflow allows surface smoothening along the envelope of the original structure. The efficient surface equilibration resulted in the transformation of a microroughness (some micrometer wide and almost 500 nm high steps) [Figs. 7(a) and 8(c)] into a nanoroughness (some 100 nm wide and less than 100 nm high steps) [Fig. 9(b)]. For the concave structures with highly varying slopes shown here, the surface smoothening seems to work over steps of up to 500 nm, from the lens bottom with the shallow ring sections up to the sharp tips with the <1 μm radius. The micrograph displaying the lens, which was obtained by using smoothening by selective thermal reflow [Fig. 9(b)], is even less rough then with the finest writing [Fig. 7(f)] and shows a superior surface finish. Although roughness values were not quantitatively analyzed using scanning force microscopy on the microlens slopes, both values displayed in Figs. 7(f) and 9(b) are well below 10 nm. This goes along with the results by other researchers for 2PP-DLW, where 10 nm roughness was obtained with 150 nm slicing with IP-S in the solid writing mode.

The sidewalls of the 3 × 3 MLA are basically the same in Figs. 8(c) and 9(c). The high aspect ratio tips were also well preserved as well as the radius of the lens curvature was maintained. This will keep the required optical diffusing functionality while homogeneity of the illumination area is improved due to less scattering from surface roughness. Small differences, e.g., the tip at the intersection of the four lenses and the right slope of the surface smoothened structure in Fig. 9(c) indicates some minor issues. It has to be mentioned that the roughness reduction by surface reflow is physically different from the writing with enlarged blurred voxels. In comparison to the enhanced overlapping of voxels, which leads to a merging of cylinder-like trajectories into a continuous surface, reflow is based on material displacement from elevated areas into small depressions and thus is a true equilibration process, which self-minimizes the total system energy. However, one should not forget that the damage of the PMMA during exposure results in a surface skin with low $M_w$, which has to be taken into account for further process steps involving temperature treatment or aggressive solvents.

5 Conclusion
The aim of this publication was to demonstrate the efforts necessary to achieve similar surface roughness by engineering process- and design-based strategies as compared with postprocessing. In particular, we compared the effect of surface smoothening due to the reduction of the slicing distance with TASTE postprocessing. For the complex concave structures shown here, similar roughness below 10-nm RMS was achieved for a slicing distance of 50 nm and the thermal reflow of original structures written with a slicing distance of 500 nm. The postprocessing method can also be tuned to modify much smaller structures than presented here using other exposure sources and photons of different wavelengths.
Independent of the speed of the original patterning, post-processing allows additional significant time reduction for the production of smooth micro-optics, which is of high industrial relevance.

The reduction of the slicing distance as well as strategies to write structural shells with smaller increment than the inner part or even a scaffolding of the lens is the way to go if high throughput structures with resolution elements have to be directly written. The postprocessing achieves low surface roughness from structures, which were optimized for speed and not for surface roughness during origination with 2PP-DLW. In the case of the MLA presented in this work, the speed would be improved about 10-fold. The trade-off is that two additional replication steps are needed. Due to the lack of availability of production machines for 2PP-DLW (apart from the newer developments presented by Multiphoton Optics GmbH) and new resists with much lower exposure thresholds, throughput will remain an issue for the fabrication of extended arrays of lenses. Scaling larger arrays and higher structures will critically increase the writing time and the postprocessing in this area will remain as a valuable and flexible method to obtain smooth surfaces. Since in the future there will be both a need for original structures and such ones being used as originals for replication, different strategies need to be developed. To judge quantitatively the structural fidelity and the deviation from the original design, for all processes, the influence of voxel size variation, slicing, process shrinkage, material loss, and displacement need to be taken into account, along with the ability to further process the structures into molds for HVM. The NIL process chain with its main elements origination, replication, and pattern transfer (functionalization) is a good example to demonstrate this capability.

Acknowledgments
This project was funded by the Swiss Nanoscience Institute (project A10.13 SurfFlow). The authors express their special thanks to K. Vogelsang (PSI) for support with the hot embossing. The authors also thank M. Rossi (Heptagon) for the support of this project and the useful discussion.

References

Robert Kirchner received his PhD in electrical engineering from the Technische Universität (TU) Dresden, Germany, in 2011. He worked as a postdoc at the TU Dresden in 2012 and at the Paul Scherrer Institute, Switzerland, from 2012 until 2017. Since 2017, he has held a DFG Heisenberg Fellowship and is leading the 3-D Mesoscopic Systems Group at TU Dresden. His current research focus is on micro-nano-origami methods, self-assembly, and advanced 3-D patterning methods.

Nachiappan Chidambaram obtained his PhD in Ecole polytechnique fédérale de Lausanne in 2014. He has worked as a postdoc at the Paul Scherrer Institute, Switzerland, since 2015. His research interests are piezoelectric materials and polymers for innovative device concepts.

Helmut Schift obtained his diploma in electrical engineering and later his PhD in mechanical engineering from the University of Karlsruhe, Germany, in 1994, following a cooperation with the Institute of Microtechnology Mainz (ICT-IMM). Since 1994, he is a staff scientist at the Paul Scherrer Institute, Switzerland, and head of the Polymer Nanotechnology Group (INKA-PSI). In 2011, he was a visiting professor at the Technical University of Denmark (DTU). He is a recognized pioneer of nanoimprint lithography and currently focuses on innovative polymer processing technologies.