Auto-Tilt Photography

Filip Sadlo[1] and Carsten Dachsbacher[2]

[1]Visualization Research Center (VISUS), University of Stuttgart, Germany
[2]Computer Graphics Group, Karlsruhe Institute of Technology, Germany

---

Abstract
Tilt-shift camera lenses are a powerful artistic tool to achieve effects like selective focus with very shallow depth of field. Typically they are used by professional photographers only, which is due to the high cost and weight, and the intricate, non-intuitive handling. We introduce the auto-tilt mechanism which is as easy to use as the standard autofocus. It allows automatic sharp focus of objects not parallel to the image plane, such as in landscape photography where getting everything sharp is often desirable. In contrast to pure computational approaches that are based on resampling from focal stacks, our approach based on true exposures enables time-dependent scenes and higher image quality. Auto-tilt can also be controlled via a simple sketching user-interface letting the photographer quickly define image regions inside and outside sharp focus. We demonstrate auto-tilt using a simple rapidly prototyped experimental setup that tilts the sensor (as opposed to classic tilt-shift lenses), and describe possible implementations in off-the-shelf cameras. We also outline future prospects with flexible image sensors currently being developed.

Categories and Subject Descriptors (according to ACM CCS): I.4.9 [Image Processing and Computer Vision]: Applications— I.3.3 [Computer Graphics]: Picture/Image Generation— I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction Techniques

---

1. Introduction

The ubiquity of digital cameras began with the availability of low priced and increasingly improved devices, and led to a flood of images and photo management and sharing applications such as Flickr. However, commodity auto-focus or single-lens reflex cameras sometimes do not provide the desirable flexibility for a certain kind of photographs. In particular the use of tilt-shift camera lenses provides artistic freedom to the photographer giving the possibility to move the lens parallel to the image plane (called shift) and thus changing the line of sight while being able to avoid the convergence of parallel object features. Using such optics also allows the rotation or swinging of the lens plane relative to the image plane (called tilt) to control the orientation of the plane of focus, which is explained by the Scheimpflug principle. In many cases, tilt-shift photography refers to the use of tilt combined with large apertures or zoom lenses to achieve a very shallow depth of field.

In this paper we introduce the auto-tilt mechanism which is as easy to use as the standard autofocus and can be used to enhance the overall sharpness in a picture, or controlled via a simple sketching user-interface for selective focus. Such cameras use an image sensor that can be tilted to exploit the Scheimpflug principle. We demonstrate the auto-tilt using a simple rapidly prototyped experimental setup and outline possible implementations in cameras.

The remainder of this paper is organized as follows: in the next section we outline previous work, and introduce the experimental auto-tilt camera, as well as its functioning, in Section 3. Section 4 presents applications followed by an outlook on possibilities with flexible image sensors (Section 5), and results in Section 6. Hardware implementation in future off-the-shelf cameras is discussed in Section 7.

2. Previous Work

Computational photography is an emerging field that strives for unbinding digital photography from being just electronically implemented film photography, i.e., from taking pictures as in the last century. As such it tries to capture information beyond just a simple set of pixels, removing constraints on dynamic range [DM97, KUWS03], depth of field [LHG*09], or motion [RAT06]. Due to the vast amount of research in recent years, this section can only give a very brief overview over research directions in this field. For a more comprehensive overview we direct the reader to the STAR [RTM*06] and survey [Lev06].

Some works in this field apply off-the-shelf cameras directly, but often the novel prospects go along with modifications to the hardware. The ultimate goal is to capture

*F. Sadlo & C. Dachsbacher / Auto-Tilt Photography*

a complete 4D light field [LH96] in high spatial and angular resolution. Existing light field cameras use lens arrays or attenuating masks trading spatial for angular resolution [Ng05, NLB*05, GZC*06, GIB07, VRA*07]. Multi-aperture photography [GSMD07] captures and manipulates multiple images of a scene taken with different aperture settings at once relying on a custom-built optical system. Using a (programmable) coded aperture [LFDF07, LLW*08] allows the modification of the frequency characteristics of defocus blur and depth estimation and high-quality light field acquisition. [BCN08] present a modified camera lens with RGB color filters placed in the aperture to capture three shifted views of a scene. This allows depth estimation and the creation of a alpha matte for a foreground object.

Several of the aforementioned ideas open up intriguing possibilities for future camera hardware if they can be built into SLR camera bodies, or even into compact cameras without negatively influencing the form-factor. Although our prototype is bulky, our method also belongs to this group, and a hardware implementation can be realized with very little space requirements. Although our method is positioned in the field of computational photography, its goals are not to enhance the possibilities of digital cameras, e.g., by capturing the entire light field, but it employs computation and interaction techniques to provide a tool to the professional as well as casual photographer.

2.1. Tilt-Shift Photography

Although pinhole cameras produce sharp images (neglecting the diffraction limit) all commodity cameras are equipped with lenses to gather more light for the film or image sensor. The main drawback of using lenses is that a limited depth of field, depending on aperture and focusing distance, is inherent. Note, that the depth of field is almost infinite for wide angle lenses and thus the tilting effect in tilt-shift photography is often used for macro photography and with zoom lenses. In a regular camera the image (or film/sensor) plane, lens plane, and plane of focus are parallel, and object regions in sharp focus are all at the same distance from the camera. The Scheimpflug principle is a geometric rule that describes the orientation of the plane of focus when the lens plane is tilted relative to the image plane (see Fig. 1). In this case all three planes intersect at a common line, and scene parts lying on a plane, but at different distances from the camera, can be brought in sharp focus. Stroebel's [Str99] excellent textbook gives an thorough introduction to the field of cameras.

Tilt-shift photography refers to two different types of movements (Fig. 2). First the inclination of the lens plane, called tilt, and second, a movement of the lens parallel to the image plane, called shift. Tilt makes use of the Scheimpflug principle and is used to control the orientation of the plane of focus, while shift changes the line of sight while keeping the image or focus plane parallel to an object. The prime example for shift is to photograph a tall building with a cam-
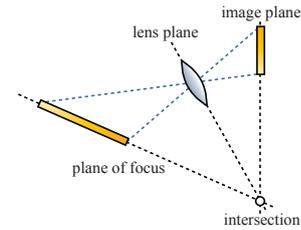
Figure 1: *Scheimpflug principle explains orientation of plane of focus when image and lens plane are not parallel.*

era pointing upwards while keeping the sides of the building parallel. Note that although most tilt-shift lenses keep the image plane fixed (with respect to the camera body) and allow the photographer to control the lens, a movement of the image plane, i.e., the image sensor, achieves the same effects. In many practical settings of photography (regarding focal length, distance to the object etc.), a comparably small inclination angle of the sensor is sufficient to achieve the desired inclination of the plane of focus.

Professional photographers either employ lens tilting for putting a certain part of the subject in focus, or for taking artistic pictures with selective focus, e.g., to direct the viewer's attention to a certain part of the image, while deemphasizing others. Selective focus is also often used to fake miniature scenes, although the effect is not exactly the same as a shallow depth of field in close-up photography.

Image manipulation software can also be used to post-process images, faking perspective and depth of field effects. However, information that is not recorded can obviously not be recovered and thus not all effects can be reproduced. Most closely related to our work is a patent describing a camera with an image sensor with five degrees of freedom for automatic shifting (for taking pictures of vertical objects) and tilting for increasing sharpness [Mut00]. The applications described therein are similar to two of ours outlined in Sections 4.1 and 4.4. Please note that our auto-tilt camera differs in several aspects. First of all, the hardware design is geared towards a realizable concept without intertial and swinging masses, and faster sensor movement. Second, although the patent outlines possible applications, no algorithm is described that is practical. Another closely related patent is [Woe09], but again no algorithms are described. Other related patents are [Sho06] and [TH06] that, however, only provide shift and tilt (swing) functionality and hence are not capable of acquiring image stacks (see Sect. 4.2 or [ADA*04]); none of the patents takes flexible sensor (or optics) into account. Other approaches, such

Figure 2: *A schematic view of a classic bellows camera with tilt and shift movements of the lens.*
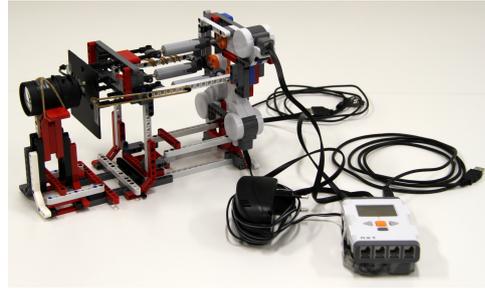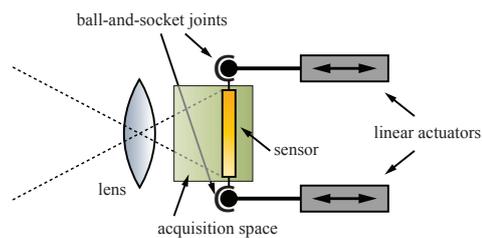
Figure 3: *Left: a schematic 2D view of the auto-tilt camera. The image sensor (yellow) is mounted on two axes controlled by linear actuators (gray) and can be moved and tilted, taking images in the acquisition space (green). Right: our experimental setup using LEGO Mindstorms.*

as [ADA*04, KNZN11], acquire image stacks and are hence not capable of acquiring single-shot images, necessary for dynamics scenes.

3. The Auto-Tilt Camera

Photography is known to be a tedious and demanding procedure. On the one hand the photographer wants to capture a moment or artistic expression, on the other hand he or she is concerned with the optimization of a multitude of parameters. As such there are view, distance and focal length, focus together with shutter and light, sensitivity of the sensor material, and its shift and, the main topic of this paper, tilt. Cameras that do not feature tilt-shift functionality already impose a hard task on the photographer regarding the choice and optimization of all these parameters. But when tilt and shift come into play, the spectrum of possibilities—but also that of difficulties—augments substantially. This paper addresses exactly this topic with the aim of supporting the professional photographer, and inviting and guiding the amateur.

A common feature in the medium and upper price segment in photography are mechanisms for shifting the image sensor or the optics laterally to compensate for undeliberate perturbations. Although they are usually limited in range, it is considered straightforward effort to increase their range for allowing the shift functionality in many cases. Our work focuses on the possibilities of a tilt (and later bend) functionality, its applications, and implementation aspects.

3.1. Experimental Setup

Our experimental setup (shown in Fig. 3) consists of commodity parts and is thus easy and cheap to reproduce. The mechanical structure is built using LEGO Mindstorms NXT parts. In particular we attached the image sensor to a carrier that is connected via ball-and-socket joints to three rods. These rods are displaced by three linear actuators, each driven by a servo motor, to control the orientation and placement of the sensor. Strictly speaking the distance between the joints increases with the inclination angle of the sensor (Fig. 3 (left)) and this would need to be compensated. In fact, these differences are small for the inclination angles

that are often required in practice and are thus negligible. The Mindstorms NXT interface allows us to control the motors via USB and Bluetooth. The image sensor is a 2 mega pixel CCD sensor (1/3.2" size) dismantled from a commodity web cam and attached to the carrier such that the centroid of the ball joints is located at the center of the CCD. The image data is transferred in raw format via USB to our application. The camera lens is a Tamron 13VM550ASII F/1.4 for 1/3" image sensors. The entire apparatus is covered using a light impenetrable box and a bellow to close the gap between the lens and sensor platform (both not shown on the picture).

3.2. Controlling the Camera Sensor

LEGO Mindstorms NXT is a system that allows the construction of robotics applications in an easy modular way. A control unit reads a multitude of sensors and drives actuators, and custom control programs can be uploaded via USB or Bluetooth. We use "Not eXactly C" (NXC), a programming language similar to C and the "bricxcc" compiler under Linux. LEGO Mindstorms' servo motors feature rotation control with an accuracy of +/- one degree according to the manufacturer. However, we experienced mainly two issues. First, we observed mechanical clearance, which is hard to cope with but it turned out to be negligible in our implementation. Second and more severe, they seem to operate at the aimed precision only during a single execution of a program, meaning that the state of the motors is not kept between program invocations. This seems negligible at first sight, however, it is likely to cause drift between successive executions because the motors are typically not actuated precisely to the destined angle. We address these issues by controlling the unit using a single program that communicates with the computer, and stores its state persistently.

The linear actuators were calibrated by manual measurement of the maximum displacement, providing control in absolute "CCD coordinates" at an accuracy better than $0.05mm$ in our setup. The location of the three rods was determined by manual measurement of the attachment points at the CCD carrier. We use two ways for defining the position of the CCD sensor: either via the displacement of each rod or by

prescribing the displacement of the center of the CCD and using two angles to describe the inclination. For most of the applications we use the latter approach.

We take the extent of the sensor into account, as in some applications we also capture the image data by moving the sensor behind the lens, effectively sampling a 3D region called *acquisition space*. Although the accuracy of the rod displacements is lower than typical pixel sizes (2.8 μm at 1600×1200 resolution in our case), we typically observe a maximum focal range of interest of 2 mm behind the lens, resulting in a $4.48 \times 3.38 \times 2$ mm^3 voxel block at typically non-isometric $1600 \times 1200 \times 200$ resolution in stack acquisition mode (Section 4.2). We would like to emphasize that we do not rely on the properties of the lens, such as focal length and distortion. All our operations are performed behind the lens both physical and computational. Thus the tedious task of lens calibration is not required and we can use any lens (zooming or fixed). In contrast to methods reconstructing object (world space) depth from focus, we only determine the physical locations in acquisition space where sharpness is highest, which is a simple and straightforward procedure.

4. Auto-tilt Operation Modes and Applications

The auto-tilt camera can be used in various operation modes ranging from fully automatic focusing to sketching sharp regions by the user. We also outline potential applications that would not be possible using a standard tilt-shift lens.

The simplest approach is to manually control tilt and position of the sensor, e.g., by manipulation of a track ball by physical buttons or by using a touch screen. However, although this approach is somewhat more convenient than inclining lens or sensor manually, because it gives better control over the aimed degree of freedom, it does not support tilt photography. Still, this might be the preferred by experienced tilt-shift photographers in some situations. We now describe the advanced techniques that support tilt photography in various ways.

4.1. Differential Auto-Tilt

The most obvious use of the auto-tilt mechanism is to extend the autofocus by tilting and moving the sensor such that the overall sharpness of an image is maximized. In this operation mode, we start similar to a standard autofocus (AF) and move the sensor—keeping it orthogonal to the optical axis—such that a chosen subject in the scene is in focus. In this case we can use existing AF mechanisms to preselect a focal range and then auto-tilt therein. Any AF system, such as through-the-lens optical AF sensors used by most modern SLR cameras, is suitable for this task. In our experiments we either adjusted focus manually, or relied on a passive AF technique measuring contrast in the image (assuming that there is enough light for passive measurements). The intensity difference between adjacent pixels, and thus contrast in the image, naturally increases with correct image focus and

is therefore one criterion used by AF techniques. Contrast-based AF does not involve distance measurement and is generally slower than phase detection systems or active AF, but easy to integrate and thus more flexible.

The AF is a good starting point for the optimal plane of focus given that any "important" subject in the scene has been brought into focus. It is also common practice in professional tilt and shift photography to first bring a part of the subject in focus and then start tilting guided by the variation of sharpness. We mimic this approach using an optimization process that tries to increase the sharpness of the entire picture. For this we use a "trial and error" strategy: we slightly tilt the sensor, and if the sharpness increases we proceed with this inclination, and begin afresh. Note that in general, a local optimization strategy like this one cannot guarantee to find the globally best-fitting plane of focus, but it performed well in our experiments.

The optimization starts with the displacement of the three rods, $r_{i,j}$; i denotes the i-th iteration of the optimization process, and $j = 0, 1, 2$ is the rod index. If an AF system determines the initial displacements, all three are equal, i.e., $d_{0,0} = d_{0,1} = d_{0,2}$. In Section 4 we will describe applications that start the optimization with different displacements. We also define a step size for every rod, $s_{i,j}$, to generate new positions and orientations of the sensor. In every iteration of the optimization procedure we take $3^3 = 27$ images, by probing all combinations of $d_{i,j} + k \cdot s_{i,j}$, for all $j = 0, 1, 2$ and $k = -1, 0, 1$. We compute the sharpness by simply integrating the contrast of the luminance of each image. As start for the next iteration we choose the combination $d_{i,j} + k \cdot s_{i,j}$ with the highest overall sharpness. When using fast sensor tilting and sharpness measurement systems a constant step size, just larger than the precision of the actuators, would be the best choice. In our experimental setup, we start with an initially larger step size of half the focal range containing sharp objects (determined by the AF). The step size is then reduced after each iteration if $k = 0$ with $s_{i+1,j} = s_{i,j}/2$, if $\max_j s_{i,j}/2 \leq s_{i,j}$. Otherwise they remain equal with $s_{i+1,j} = s_{i,j}$. The process terminates as $\max_j s_{i,j}$ drops below mechanism precision.

4.2. Least Squares Auto-Tilt

The differential auto-tilt is the method of choice for photographs focusing on compact regions where the untilted AF gives a good initial estimate. However, in cases where the desired regions of focus are not compact and at different depth, the estimate and hence the differential refinement will usually not converge to the desired result. This motivates another approach: a global search for maximum sharpness.

This approach relies on sampling of the acquisition space, i.e., the acquisition of a stack of images at different *z*-offsets without tilt (similar to [ADA*04], see Fig. 4). Having this data, we can define the optimal position and orientation of the sensor with respect to maximum sharpness in terms of a

F. Sadlo & C. Dachsbacher / Auto-Tilt Photography
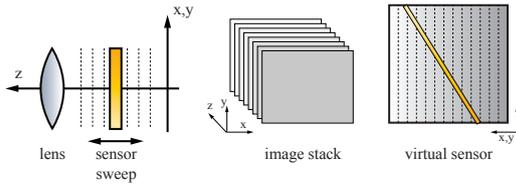
Figure 4: *By sweeping the sensor we record slices of the acquisition space. This data is used for least squares auto-tilt and to render images of arbitrarily placed virtual sensors.*

least squares problem. We first determine the z-coordinate z_i of maximum sharpness σ for each pixel of the sensor. Note that we do not reconstruct object-space depth from focus which would require tedious and lens-dependent optical calibration of the system. Identifying the positions of maximum sharpness as observations in terms of least squares, the position and orientation of the sensor can easily be determined by fitting a plane to these points.

The classical depth from focus approach is known to be a hard problem with respect to robustness, outliers are a common problem. The fact that we perform the evaluation in acquisition space is no remedy; we too have to reject outliers to obtain a representative fit. This is achieved by a simple heuristic: for each sensor pixel, we compute its sharpness in all slices of the stack, determine the z-coordinate z_i of the slice with maximum sharpness σ, and then reject the pixel if sharpness does not sufficiently satisfy monotonicity: we penalize $\partial\sigma/\partial z < 0$ for $z < z_i$, and $\partial\sigma/\partial z > 0$ for $z > z_i$. Imposing a threshold (percentile) on $\sum_z sgn(z - z_i)\partial\sigma/\partial z$ with z being the z-coordinate of the respective slice, turned out to be a robust strategy for obtaining appropriate observations (pixels) for the least squares problem; $\partial\sigma/\partial z$ is evaluated using finite differences.

In anticipation to Section 5 where we take an outlook to bendable sensors, we do not implement the least squares problem as linear regression, i.e., fitting a plane. Instead, we use the more generic model of thin plate splines (TPS) which are modeled using the basis function $\phi(r_j) = r_j^2 \log(r_j)$, with r_j being the Euclidean distance from the basis function j to the point where the spline is evaluated, and coefficients c_j. These splines have a physical meaning: they correspond to the deformation of a thin metal plate when orthogonal displacements are enforced at $n_t \geq 3$ given positions, i.e., they interpolate between these points minimizing the bending energy. If $n_t = 3$, TPS are planes and this allows us to treat rigid and flexible sensors in a single framework where we can freely choose the number n_t of displacing actuators. The basis functions are superposed by a linear function, leading to $n = n_t + 3$ unknowns: $c_1, \ldots, c_{n_t}, a_0, a_1, a_2$. There are three additional equations, hence $m = m_o + 3$ equations in total. The coefficients are typically determined by prescribing a value z_i at each of the j basis functions, which would lead to a symmetric $(m = n)$-matrix that could be directly solved. In our case, the positions $\mathbf{x}_i$ of the observations are

the m_o sharpest pixels, their depth of maximum sharpness is z_i, and the basis functions correspond to the actuators at $\mathbf{x}_j$. Hence, we cannot assume that there is a valid observation (a pixel that passed the monotonicity test) at each actuator, and typically there are very many observations. Hence the $m \times n$ overdetermined system is suited for a least squares approach:

$$
\begin{aligned}
\sum_{j=1}^{n_t} c_j \phi(\|\mathbf{x}_i - \mathbf{x}_j\|) &+ a_0 &+ a_1 x_i &+ a_2 y_i &= z_i \\
\sum_{j=1}^{n_t} c_j &+ 0 &+ 0 &+ 0 &= 0 \\
\sum_{j=1}^{n_t} c_j x_j &+ 0 &+ 0 &+ 0 &= 0 \\
\sum_{j=1}^{n_t} c_j y_j &+ 0 &+ 0 &+ 0 &= 0.
\end{aligned} \quad (1)
$$

Since the last three equations represent hard constraints, we apply weighted least squares and assign a very high weight to these. We did this for convenience, a cleaner approach would be to incorporate these constraints into the least squares solver.

4.3. Virtual Sensor Trackball

This operation mode involves user interaction to adjust the plane of focus. At the beginning, we sample the acquisition space from minimal to maximal displacement, taking images at equidistant steps (typically between 30 and 100 images). We treat the image stack (Fig. 4) as a 3D texture where the x and y axes span the image plane, and the z-axis is pointing in the sweep direction. This allows us to quickly render preview images of arbitrarily placed sensors as the images in the stack represent slices of the acquisition space, i.e., the volume of locations where light can be measured by different sensor placements (Fig. 3 (left)). We allow the user to position the virtual sensor using a trackball plus displacement control and then sample the respective slice of the 3D texture. In an interactive application (Fig. 7 (top right)) we use OpenGL to render a quadrilateral covering the entire screen. The 3D texture coordinate for every pixel is computed by intersecting a ray in z-direction through that pixel with the virtual sensor. The missing information between two slices is tri-linearly interpolated by the texturing hardware. Once the user has adjusted the virtual sensor placement, we can position the real sensor of the auto-tilt camera accordingly, and record a high-quality image. Note that given a reasonably high number of slices, the quality of the image obtained from interpolating the slices is of good quality, although supposedly not satisfactory for the professional photographer. Further, our approach using the real sensor can acquire dynamic scenes with a single shot, whereas instantaneous stack acquisition is not feasible with today's hardware.

There are several applications that benefit from this operation mode. One example is power-critical photograph: depending on the hardware implementation it might be too power consuming to tilt the sensor while the photographer investigates the best acquisition. With flash tilt photography in the dark it is also not desirable to continuously flash the scene. Acquiring the stack once and investigating the optimal setting virtually is an attractive alternative in these scenarios. However, the virtual trackball requires using a tripod

or to use image registration to make sure that the parameters obtained virtually also render the desired results later in the real photograph. Note that there are situations where the usage of a tripod is dispensable, e.g., if the differential method is issued right before the final photograph and the viewpoint and parameters of the scene did not change substantially. Further applications include scenes that vary slowly enough to acquire them in a stack, but too fast to be able to investigate the best artistic tilt configuration. Lastly, the virtual tilting can provide images at extreme inclinations, and hence easing the mechanical design of the tilt mechanism by limiting the physical inclination.

4.4. Sketching Sharpness

Tilt-shift photography is often used as an artistic tool to put some objects into, and others out of focus. We provide a simple to use sketching interface for the photographer to achieve this effect without further manual interaction. Again, we record the image stack and display the image with a planar least squares fit to provide a reasonably clear impression of the scene. Next, we let the user mark regions of the image that should be in sharp focus using few strokes (Fig. 5 and 7). In our implementation this is done using the mouse or a touch pad, in a real camera this can easily be achieved by providing a touch screen. Then we can either use the differential auto-tilt or the least squares fit, both for planar and bent sensors (Sect. 5), to obtain the desired image. The marked pixels are used to control and weight the fitting procedures, enforcing sharpness for the respective image regions. We can either reconstruct the final image from the image stack, or position the real sensor to take a high-quality photograph.

5. Auto-Bend Photography

In theory we can modify the virtual sensor application easily such that we render an image of maximum sharpness: instead of placing a *planar* virtual sensor we sample the slice providing sharp focus for each image region or pixel. This could be easily achieved by precomputing the sharpness for every slice in the 3D texture, and when computing the texture coordinate for a pixel (x, y) we search along each ray $r(t)$ for maximum sharpness. This would result in arbitrarily shaped virtual sensors, however, any rendering using this method would be error prone and would inevitably lead to artifacts in the image, e.g., at depth discontinuities.

Although increasing the overall sharpness of photographs is certainly an appealing goal we make two observations: first, a human observer is used to a certain degree of defocus in pictures and movies. The defocus is due to the limitations of traditional photography on the one hand, but also intentionally used by photographers to steer the observer's attention. Note that in both cases the defocus strengthens the depth perception and we believe that the possibility of this visual cue should be preserved in photography. The second observation is that great progress is currently being made in

the field of fabricating flexible polymer-based light sensor arrays [NWL*07]. This allows us to design cameras whose sensor is not only inclinable, but also bendable. Following these ideas we propose auto-bend photography which can be implemented using such flexible image sensors to provide additional freedom for controlling the—no longer planar—surface of focus. Obviously there is no such high-quality sensor available nowadays, but we can simulate the result images using a recorded image stack.

The implementation of a preview rendering of auto-bend photography is straightforward. Instead of intersecting the view rays with a plane (as in Section 4.3) we evaluate the thin plate spline at the respective location to obtain the z-texture coordinate (Section 4.2). Fig. 6 and 7 show two examples how auto-bend photography might look like (virtual photographs are indicated by a window border).

6. Results

Tilt photography is especially valuable in the field of macro photography, and we restrict the examples in this section to this field. We used our prototype setup and the described techniques for three different examples, each demonstrating maximum benefit from a different variant of our method.

Coin This example of a coin on a sheet of uncoated paper (Fig. 5) is a prominent case for tilt photography. It illustrates the result using autofocus, which only puts a small part into focus, and plane-fitting which improves the overall sharpness. Unfortunately, the result is suboptimal mainly due to one reason: the paper features sufficient structure to cause substantial sharpness and hence gaining importance with regard to auto-tilt focus. The sketch-based plane fit solves this problem lifting the plane of focus to the coin.

Paperweight The second example treats a difficult case in tilt-macro photography: a sphere. Even more tricky, it is made of glass with highly reflective particles below the surface. Fig. 6 shows pictures taken with standard autofocus, using auto-tilt, and virtual auto-bend photography with 3×3 uniformly distributed rods for the TPS. It can be seen that the region in focus extends almost to the silhouette.

Game pieces Fig. 7 shows a setup of game pieces and pictures taken with autofocus, least squares auto-tilt, and sketch-based photography with planar and bent sensors. Auto-bend photography, again with 3×3 uniformly distributed rods, yields very good results; the dog's eyebrows are in perfect focus at the cost that other regions slightly lose sharpness. This is also an example that TPS tends to oscillate. Although it is physically-based and minimizes bending energy, there are possibly interpolation functions that could be better suited for this task, especially in virtual auto-bend where the physics of the sensor does not play a role.

7. Hardware Implementation

It is relatively simple to motorize existing tilt-shift cameras (or lenses) to incorporate the proposed auto-tilt functionali-

Figure 5: *From left to right: 1) a photograph of a coin using standard autofocus (only a small part is in focus). 2) fitting a plane balances the sharpness (or unsharpness) between the coin and the structured surface below. 3) using the sketching interface we mark to coin to bring it into focus. 4) the coin and all surface features are in focus, the ground plane is slightly blurred compared to image 2. That is, the focus plane is now well fit to the coin's upper surface (please zoom in).*

Figure 6: *Please zoom the electronic version of this paper to better notice the differences in the photographs. Left: a picture taken using a standard autofocus (note the sharp center, and unsharp boundary area). Center: using a plane fit, the plane of focus is slightly further away from the camera, and thus slightly unsharp regions appear at the center, and again at the boundary. Right: using a (virtual) bent sensor controlled by 3×3 rods it is possible to put almost the entire sphere into focus.*

Figure 7: *Top-left: a photograph of game pieces arranged in an arc using standard autofocus. Top-center: unconstrained fitting of a plane yields almost acceptable results, however, especially the dog in the left front is in defocus. Top-right: a synthetic photograph with an artistic plane of focus generated from the image stack with our application. Bottom-left: we use the sketching interface to mark the game pieces as most relevant. Bottom-center: the weighted plane fitting, according to the sketched regions, yields much better results, yet not all pieces are in perfect focus. Bottom-right: Using a bent sensor (simulated in our application for 3×3 rods), we produce a photograph with all pieces rendered sharp.*

F. Sadlo & C. Dachsbacher / Auto-Tilt Photography

ties. However, it is less evident how to make this technique available in commodity cameras with as little effort as possible. Many camera manufacturers feature models that compensate for both angular and lateral camera shake. This can be achieved by moving the sensor or adapting the lens. There has been significant progress over the last decade in the way how lenses or sensors are actuated. Whereas camera shake plays little role at short focal lengths, it becomes an issue at long lengths and, in particular, in macro photography. This is the reason for the increasing interest of the manufacturers in coming up with image stabilizers for macro photography. This may result in a synergy with tilt-shift photography. Although image stabilizers compensating angular shake by actuating the lenses are designed not to effect tilt because this would lead to variation in focus, they can serve as a basis for the development of auto-tilt functionality. Alternatives for flexible sensors to achieve the auto-bend functionality can be found in the field of adaptive optics. Adaptive mirrors are a well-studied and advancing technology and offer interesting possibilities in combination with auto-tilt of rigid sensors.

8. Conclusions

In this paper we demonstrated the auto-tilt mechanism and its prospects by means of a rapidly prototyped experimental setup. We outlined several application scenarios and showed result images. We also introduced auto-bend photography which describes the possibilities of flexible image sensors that are currently being developed. A major advantage of our technique compared to image stack based synthesis approaches is the ability to capture true instantaneous result photographs, allowing for dynamic scenes.

Acknowledgements

References

[ADA*04] AGARWALA A., DONTCHEVA M., AGRAWALA M., DRUCKER S., COLBURN A., CURLESS B., SALESIN D., COHEN M.: Interactive digital photomontage. *ACM Trans. Graph. (Proc. of SIGGRAPH) 23* (2004), 294–302. 2, 3, 4

[BCN08] BANDO Y., CHEN B.-Y., NISHITA T.: Extracting depth and matte using a color-filtered aperture. *ACM Trans. Graph. (Proc. of SIGGRAPH Asia) 27*, 5 (2008), 134:1–134:9. 2

[DM97] DEBEVEC P. E., MALIK J.: Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH '97* (1997), pp. 369–378. 1

[GIB07] GEORGIEV T., INTWALA C., BABACAN D.: Light-field capture by multiplexing in the frequency domain. Technical Report, Adobe Systems Inc., 2007. 2

[GSMD07] GREEN P., SUN W., MATUSIK W., DURAND F.: Multi-aperture photography. *ACM Trans. Graph. (Proc. of SIGGRAPH) 27*, 3 (2007). 2

[GZC*06] GEORGEIV T., ZHENG K. C., CURLESS B., SALESIN D., NAYAR S., INTWALA C.: Spatio-angular resolution tradeoff in integral photography. In *Proc. of Eurographics Symposium on Rendering* (2006), pp. 263–272. 2

[KNZN11] KUTHIRUMMAL S., NAGAHARA H., ZHOU C., NAYAR S. K.: Flexible depth of field photography. *IEEE Transactions on Pattern Analysis and Machine Intelligence 33* (2011), 58–71. 3

[KUWS03] KANG S. B., UYTTENDAELE M., WINDER S., SZELISKI R.: High dynamic range video. *ACM Trans. Graph. (Proc. of SIGGRAPH) 22*, 3 (2003), 319–325. 1

[Lev06] LEVOY M.: Light fields and computational imaging. *Computer 39*, 8 (2006), 46–55. 1

[LFDF07] LEVIN A., FERGUS R., DURAND F., FREEMAN W. T.: Image and depth from a conventional camera with a coded aperture. *ACM Trans. Graph. (Proc. of SIGGRAPH) 26*, 3 (2007), 70. 2

[LH96] LEVOY M., HANRAHAN P.: Light field rendering. In *SIGGRAPH '96* (1996), pp. 31–42. 2

[LHG*09] LEVIN A., HASINOFF S. W., GREEN P., DURAND F., FREEMAN W. T.: 4d frequency analysis of computational cameras for depth of field extension. *ACM Trans. Graph. (Proc. of SIGGRAPH) 28*, 3 (2009), 97:1–97:14. 1

[LLW*08] LIANG C.-K., LIN T.-H., WONG B.-Y., LIU C., CHEN H. H.: Programmable aperture photography: multiplexed light field acquisition. *ACM Trans. Graph. (Proc. of SIGGRAPH) 27*, 3 (2008), 55:1–55:10. 2

[Mut00] MUTZE U.: Electronic camera for the realization of the imaging properties of a studio bellow camera. US Patent 6072529, 2000. 2

[Ng05] NG R.: Fourier slice photography. *ACM Trans. Graph. (Proc. of SIGGRAPH) 24*, 3 (2005), 735–744. 2

[NLB*05] NG R., LEVOY M., BRÉDIF M., DUVAL G., HOROWITZ M., HANRAHAN P.: Light field photography with a hand-held plenoptic camera. Stanford University Computer Science Tech Report CSTR 2005-02, 2005. 2

[NWL*07] NG T., WONG W. S., LUJAN R. A., APTE R. B., CHABINYC M., LIMB S., STREET R. A.: Flexible, polymer-based light sensor arrays on active-matrix backplanes fabricated by digital inkjet printing. Materials Research Society Spring Meeting, San Francisco, CA, USA, 2007. 6

[RAT06] RASKAR R., AGRAWAL A., TUMBLIN J.: Coded exposure photography: motion deblurring using fluttered shutter. *ACM Trans. Graph. (Proc. of SIGGRAPH) 25*, 3 (2006), 795–804. 1

[RTM*06] RASKAR R., TUMBLIN J., MOHAN A., AGRAWAL A., LI Y.: State of the art report: Computational photography. Eurographics, 2006. 1

[Sho06] SHONO T.: Digital camera having a tilting/swinging mechanism. US Patent 7064789, 2006. 2

[Str99] STROEBEL L.: *View Camera Technique*. Focal Press; 7th edition, 1999. 2

[TH06] TAKAHASHI K., HORIUCHI A.: Image sensing system and its control method. US Patent 6985177, 2006. 2

[VRA*07] VEERARAGHAVAN A., RASKAR R., AGRAWAL A., MOHAN A., TUMBLIN J.: Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing. vol. 26, p. 69. 2

[Woe09] WOEHLER C.: Digital camera with tiltable image sensor. US Patent 6567126, 2009. 2