EUROGRAPHICS 2015 / O. Sorkine-Hornung and M. Wimmer
(Guest Editors)

Volume 34 (2015), Number 2

Light Chisel: 6DOF Pen Tracking

V. Bubník and V. Havran

Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic

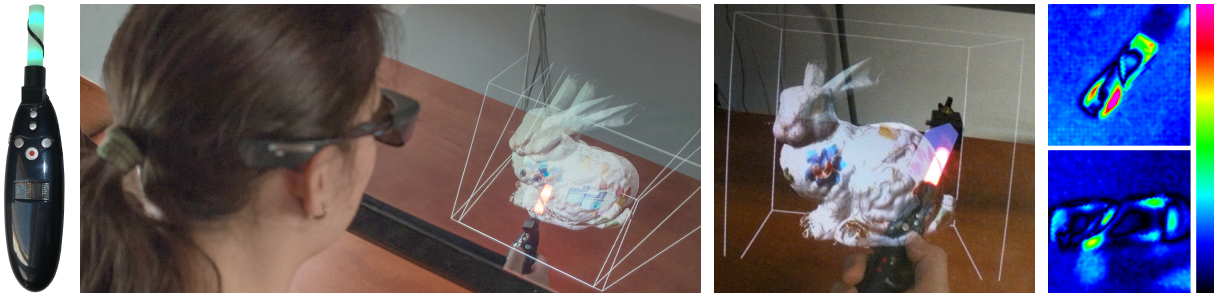


Figure 1: *The Light Chisel:* ***(left)*** *Prototype of our device.* ***(left middle)*** *User interaction with the Light Chisel in an augmented reality setup,* ***(right middle)*** *a close-up of our direct 3D modeling application work space, and* ***(right)*** *improvement of the 6DOF pose by template matching, showing the difference between the Light Chisel diffuse cylinder model and the physical tip projection in a camera image.*

Abstract

We present a novel interaction device tracked in 6 degrees of freedom by two commodity cameras. The inexpensive Light Chisel is statically illuminated with two LEDs, and uses no additional sensor (e.g. inertial or magnetic) or means of communication or synchronization. Its form factor is well suited for a screwdriver or chisel grip, allowing the Light Chisel to be rolled between the fingers. The position and orientation of the tool is tracked absolutely, making the Light Chisel suited for complex interaction, e.g. geometric modeling in augmented reality. The Light Chisel is physically small, limiting the physical and optical collisions with the real world. The orientation of the tool is tracked in a wide range of angles: pitch and yaw $\pm 90^\circ$*, roll* $\pm 180^\circ$*. We evaluated our system against the OptiTrack optical tracking system. Our system achieved mean differences from OptiTrack reference of* 2.07 mm *in position,* 1.06° *in yaw and pitch, and* 5.26° *in roll using a pair of VGA cameras. We demonstrate usefulness of our Light Chisel in four applications: character animation, modeling by swirls, volumetric modeling, and docking of CAD models.*

Categories and Subject Descriptors (according to ACM CCS): I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Tracking H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities H.5.2 [Information Interfaces and Presentation]: User Interfaces—Input devices and strategies

1. Introduction

Recent widespread use of 3D displays permitting virtual and augmented reality applications on consumer level hardware together with the increase in the computational power of GPUs has enabled modes of user interaction that were not possible before. We introduce here a cheap but powerful interaction device suitable not only for geometric modeling but also for other types of interaction with computer graphics content. The device allows for six degrees of freedom (6DOF) tracking, three for positions, two for orientation (pitch and yaw), and one for roll. The tool shown in

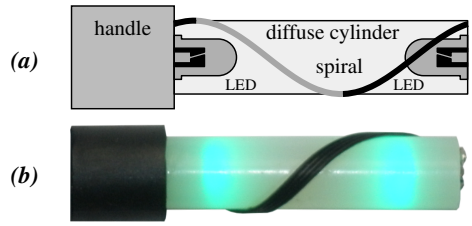


Figure 2: *Tip of the Light Chisel:* ***(a)*** *Tip geometry,* ***(b)*** *Photograph of the Light Chisel tip, light on.*

Figure 1 (left) is called the *Light Chisel*, it requires a novel optical tracking algorithm, which we present in the paper.

We believe, that the advent of consumer 3D capable displays together with the direct 3D user input devices and 3D scanners has a potential to introduce the 3D content creation to the masses. Often the steep learning curve of a mouse driven navigation in 3D world hinders creativity or is a showstopper to non-technical people. In contrast, even a child is able to sculpt in 3D using our Light Chisel in an AR environment (see Figure 1 middle). We provide the user with a convenient 6DOF tool for navigation, component selection and placement, character animation, digital painting, sculpting, material layering or gaming. 6DOFs are required by the chisel metaphor for sculpting. Our tool is uniquely suited for, but not limited to a table top application, where common optical tracking techniques are difficult to apply. A typical 6DOF tracking target is bulky. It carries 4 to 6 retro-reflexive balls, which collide with the physical environment and suffer from optical occlusions by hands. Only a large number of cameras is able to deal with the occlusions.

Our Light Chisel device consists of a statically illuminated diffuse cylinder attached to a handle, see Figure 1 (left). Our Light Chisel prototype is 195 mm long and weighs 67 g including batteries. The tool is thin and lightweight enough to make rolling inside the palm easy, utilizing the full 360° of roll. Only two cameras are necessary to track the pitch and yaw of the tool in a range of $\pm 90°$ without occluding the illuminated cylinder by hand, independent of the grip (pen grip, screwdriver grip or inside the palm). We use a diffuse cylinder 42 mm in length and 10 mm in diameter for our table-top geometric modeling application, but the cylinder up-scales naturally to the CAVE environment or to Nintendo WiiMote like computer games.

Figure 2 shows the Light Chisel tip in detail. The diffuse cylinder is illuminated with two RGB LEDs, and it is augmented with a single turn of a spiral on its surface, allowing detection of the roll angle, while the push buttons on the tool handle change the emitted color. The use of commodity RGB cameras makes optical decoding of the commands possible. Color LEDs emit light with high saturation, permitting segmentation by hue, thus improving Light Chisel detection robustness in the presence of light spots in the camera view. A prototype of the device in its basic form can be built for less than $10 for the parts.

The illuminated diffuse cylinder offers the following benefits: 1) The Light Chisel tip is large and bright enough to cause a distinct response of a computationally efficient blob detector, allowing real time tracking by limiting further processing to a small number of image regions. Nevertheless, the light energy is diffused to a large enough volume not to blind the user. 2) The axis of the cylinder projection in the camera image is calculated by principal component analysis quickly and with high accuracy. 3) We have found a simple yet sufficient model of the diffuse cylinder, which we project to a camera image for template matching to improve the final 6DOF pose of the diffuse cylinder, see Figure 1 (right) for an example. 4) No expensive external illumination is required as in the case of passive retro-reflexive markers.

Our contribution consists of (1) the design of a simple cylindrical light emitting object, and (2) the corresponding interactive algorithm to track the object in 6DOF by off-the-shelf cameras with sufficient precision and accuracy.

The paper is further structured as follows: In section 2 we survey the related work, in section 3 we present our Light Chisel design and the tracking method, in section 4 we present the measurement results, and in section 5 we conclude the paper, together with hints for future work.

2. Related Work

Various physical principles are used to capture the absolute tool pose, the most common being optical, inertial, magnetic, ultrasound or mechanical. Hightower and Borrielo [HB01] and Welch and Foxlin [WF02] give an extensive overview of 3D tracking techniques. Many of these have been commercialized for motion tracking in the movie industry or in computer-assisted surgery. In this work, we are interested in systems capable of 6DOF tracking.

Optical tracking, camera based: Passive retro-reflective balls or active IR markers are tracked with multiple IR cameras. Passive retro-reflective markers are popular because of their low price point. Companies like OptiTrack, ARTrack, iotracker and NDI provide various forms of tracking targets. Passive targets expose a number of retro-reflective balls in a configuration, which allows for unique 6DOF pose estimation. Pintaric and Kaufmann [PK08] discuss passive target design. While the targets offer good accuracy, they are bulky. For example, the 6 reflective balls of the ARTrack Flystick target occupy the space of a ball 170 mm in diameter. Even illumination of the tracking space is required to correctly detect centroids of the retro-reflexive balls. For example, each OptiTrack S250e camera [Opt14] carries cca. 100 IR LEDs. On the contrary, our Light Chisel requires two LEDs only. Active IR markers are time multiplexed, requiring fast cameras to maintain a high data acquisition rate.

A Bokode [MWH*09] device consists of a microfilm with a tiny lens, illuminated by a diffuse LED. Together with a camera lens, the Bokode integrated lens act as a microscope magnifying a tiny spot of a microfilm to the camera sensor. The position of the Bokode emitter in relation to the camera is calculated from the perspective distortion of a Bokode pattern. The orientation is calculated from the part of the microfilm that is captured. A Bokode emitter is visible in a limited angular range ($\pm 20^\circ$) towards the camera axis, and it requires large aperture sensitive camera and long exposition times, leading to motion blur.

Optical Tweezers [MG03] and Sceptre [WNG*06] calculate up to a 6DOF pose of a modified IR laser pointer emitting a fixed pattern onto a projection screen, which is captured from the rear side by an IR camera. The range of orientations where the projected pattern hits the screen is limited, so multiple emitters or screens are required to cover a wider range of orientations.

Wang and Popović [WP09] track hands wearing color patched gloves by template matching against an image database. Wang [WTK13] further extends the technique to depth sense cameras. Further progress on hand tracking by depth sense cameras is demonstrated by [QSW*14] and [TSLP14], but the results are difficult to interpret with regard to precise 6DOF manipulation. Wang et al. [WPP11] recommend to track only 3DOF of each hand in a CAD application to avoid excessive wrist strain. On the contrary, a physical device like our Light Chisel can be rotated and rolled between the fingers for 6DOF input with little effort. A simple physical object with known dimensions like the Light Chisel is certainly easier to be tracked with high accuracy than a highly variable human hand. The Leap Motion controller tracks hands or pen-like objects by capturing silhouettes by 2 IR cameras and 3 IR LEDs packed in a tiny table top box. Finger-like objects with an elliptical cross-section are reconstructed from silhouettes [Hol13]. The Leap controller is accurate [WBRF13, GJP*14], but finger overlaps are not handled. Therefore 5DOF is supported in a wide range, but tracking of roll is limited.

Combined optical & inertial tracking: The Nintendo Wii controller integrates an IR camera tracking a fixed static LED pattern to derive its position. The Sony PS3 Console tracks an emissive ball by a single static RGB camera. Both controllers track their orientation by inertial sensors. Sony also senses Earth magnetic field. zSpace [zSp13] equips its 3D displays with a 6DOF stylus. A single IR LED at the tip defines its position, and the orientation is tracked by inertial and magnetic sensors. The WorldWiz Wand [Wor13] carries two active optical markers for 5DOF tracking, and is equipped with inertial sensors. Integrating the readings of the differential inertial sensors leads to over-shots and absolute offsets, which are often corrected by magnetic sensors tracking the Earth magnetic field, distorted by close metal objects. In contrast to the combined optical / inertial systems, our Light Chisel tracks all 6DOF optically, therefore achieving better registration with the real world.

Optical tracking, non-camera based: Single point photo sensors, 1D CCD arrays or a 2D position sensitive photo detector provide significantly higher frame rates than CCD or CMOS cameras, allowing time or frequency multiplexing of active markers, and frequency modulation / filtering for ambient light suppression.

Welsh and Bishop [WB97] track a golf ball sized optical sensor head from time multiplexed dense LED marker panels. The expensive lateral sensitive photo detectors provide one sample per sensor at a 3kHz rate, fused by Kalman filters.

Raskar et al. [RNd*07] label space by projecting orthogonal time divisioned Gray coded stripes. A sensor reads its position off the projected Gray code synchronously with the projector. The orientation is sensed from the incident light emitted by additional time multiplexed IR LED beacons, taking advantage of the near Lambertian dependency of the LED emissivity and the photo-detector sensitivity on the off-axial angle. IRCube [HHC*11] / IRPen [HHLL14] applies the LED and photo detector angular sensitivity to 6DOF tracking of an 18 IR LED head by 4 photo detectors. The IrPen is sensitive to the calibration of the IR LEDs and sensors, which is a labour intensive process. Numerical readings of IR sensors are influenced by the temporal and thermal stability of all the components of the system (sensor sensitivity, LED angle dependent emissivity, amplifier gain, dirt on optical components), by the ambient lighting and by reflection from nearby objects. IRPen is closest to the form factor of our tool, is able to deal with limited occlusions and tracks a wide range of orientations. In contrast to our system, it requires custom electronics at both the transmitter and the receiver, requires synchronization of the hand held device, and the LED head draws twice the current of our Light Chisel.

Lumitrack [XHW*13] projects two perpendicular time multiplexed m-sequence bands onto at least three stationary 1D CCD sensors for 6DOF tracking. M-sequence is such a sequence, of which each consecutive M symbols long subsequence is unique. Using the minimal sensor and projector configuration, the tracking range of the tool orientation is limited. Kinetrack [MIF12] decouples the random dot pattern emitter from the Kinect device and tracks it in 6DOF by matching the point pattern projection onto a known room geometry. Projector-based techniques like Lumitrack or Kinectrack require high-power light sources, necessitating wired connection or heavy batteries.

Haptic tracking: Novint Falcon haptic device offers 3DOF sensing and force feedback. Geomagic Touch (former Sensable Omni) offers 6DOF sensing and 3DOF force feedback. These devices are expensive, fragile and have limited working range. Jacobson et al. [JPG*14] devised modular haptic blocks, from which a haptic proxy to a character skeleton is

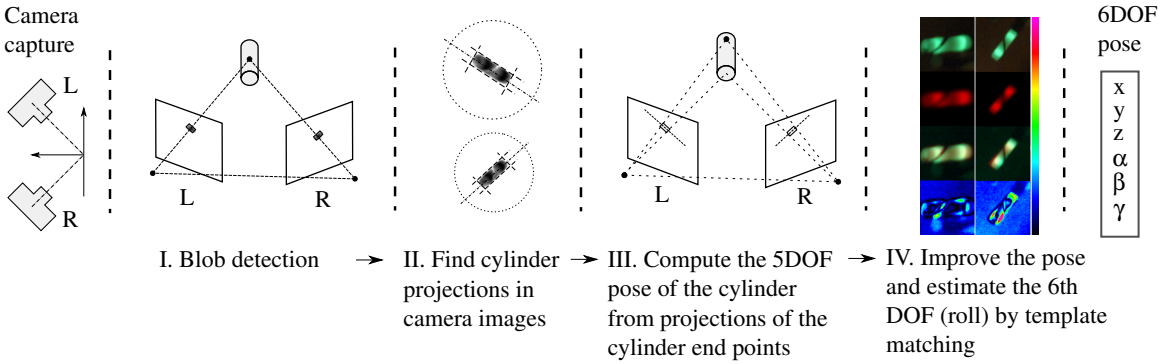


Figure 3: *Overview of the tracking algorithm pipeline consisting of four stages, from stereo camera capture to the 6DOF pose.*

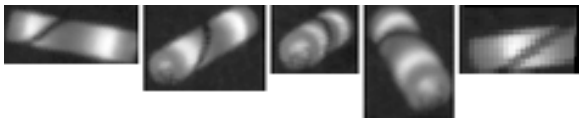


Figure 4: *Examples of Light Chisel cylinder images as seen by the cameras, interpolated from Bayer mosaic. The average length of the cylinder projection in camera is 46 pixels.*

assembled. The method has its use limited to character animation.

3. Our Approach

Below we describe our interaction tool design and then we introduce our algorithm for 6DOF tracking of our tool.

3.1. Interaction Device

The Light Chisel is formed by a short diffuse cylinder made of polyamide plastic with two RGB light emitting diodes countersunk into its sides, see Figure 2(a). Our prototype cylinder is 42 mm in length and 10 mm in diameter. There is a single turn spiral wound along the cylinder length to provide visual information on the rotation of the cylinder around its axis (roll) and to power the light emitting diode on the tip. The cylinder emits green color when no button is pressed. Pressing the buttons on the handle changes the emitted color to red, blue or any combination of the three LED chips. A change in the color when the button is pressed is detected by the tracking algorithm as a command. Figure 4 shows how commodity CMOS cameras see the diffuse cylinder with its two LEDs lit. Because the diffusion ability of the cylinder material is limited, the intensity of the emitted diffused light varies strongly and a commodity camera cannot capture both light regions without saturation and dark regions with sufficient signal to noise (SNR) simultaneously. Over-exposition leads to undesirable effects like blooming and unequal clipping in the RGB color channels, leading to a shift of hue. In the tracking algorithm, hue is used for segmentation and detection of commands, we therefore opted to avoid camera saturation, sacrificing SNR of the cylinder edge regions to keep the colors consistent.

3.2. Tracking Algorithm

The tracking algorithm inputs a pair of camera captured images and performs the following steps to extract the 6DOF pose of our Light Chisel (see Figure 3):

I. Pairs of left / right high intensity blobs are detected, satisfying the epipolar constraints [HZ03]. For each blob pair, a respective 3D point is calculated.
II. Oriented bounding boxes enclosing the Light Chisel projection are found in the left / right camera image for each blob pair. Inside the oriented bounding boxes the projections of the cylinder end points are calculated.
III. Using the left / right pair of oriented bounding boxes and cylinder end point projections in the camera image space, the 5DOF pose of the cylinder in the world coordinate system is calculated.
IV. The position / orientation is improved and the 6th DOF (rotation around the cylinder axis, roll) is recovered by template matching.

Below we elaborate on these steps in detail.

I. Find candidates for 3D cylinder positions: High intensity blobs are detected by a Laplacian of Gaussian (LoG) blob detector in the left / right camera and all combinations of left / right camera blobs are tested against the epipolar constraints, resulting in a limited number of accepted blob pairs. The LoG detector is tuned to an average width of the cylinder projection in the camera, so at least one blob is detected per cylinder projection and at least one blob pair fulfills the epipolar constraints. A blob pair is dropped if the left blob does not project on an epipolar line of the right blob in the left camera, and vice versa, see Figure 5 for an example. The cylinder projects to blobs B_{1L} and B_{1R}. Blob B_{1R} falls on an epipolar line e_{1R}, defined as an intersection of an

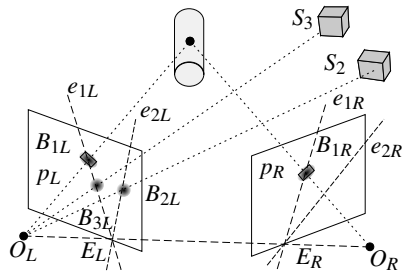


Figure 5: *Blob detection in the camera image space. Projections of spurious bodies S_2, S_3 are filtered out using epipolar constraints.*

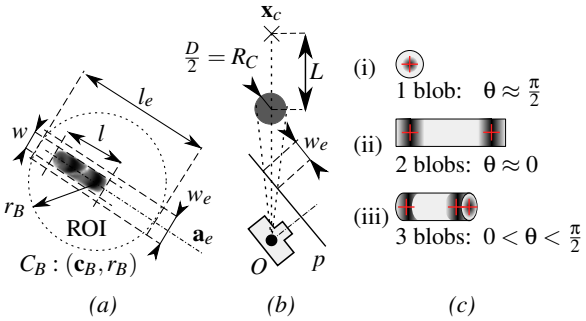


Figure 6: ***(a)*** *Detecting an oriented bounding box of a Light Chisel cylinder projected into the camera space.* ***(b)*** *Estimating the width w_e of the cylinder projection.* ***(c)*** *Blobs detected inside the axis aligned bounding box B_e, dependent on angle θ of the cylinder axis towards the camera plane.*

epipolar plane (O_L, O_R, B_{1L}) and the right camera plane p_R. Similarly, blob B_{1L} falls on an epipolar line e_{1L}, defined as an intersection of an epipolar plane (O_L, O_R, B_{1R}) and the left camera plane p_L. Blob B_{2L} represents a spurious object S_2 outside of the field of view of the right camera, so no blob is found on epipolar line e_{2R} and B_{2L} is dropped. A spurious object S_3 projects as B_{3L} on the same epipolar line e_{1L} as B_{1L}, so blob pairs (B_{1L}, B_{1R}) and (B_{3L}, B_{1R}) are considered for further processing. The invalid blob pair (B_{3L}, B_{1R}) will be filtered out in the following steps by dissimilar hue, physical cylinder length constraints or template matching quality. For each remaining blob pair (B_L, B_R), 3D position of the Light Chisel center $\mathbf{x}_c$ is estimated as the closest point of rays (O_L, B_L) and (O_R, B_R). Blob pairs with $\mathbf{x}_c$ outside of the tracking space are dropped.

II. Find cylinder projections in camera images: For each candidate 3D position $\mathbf{x}_c$, axis $\mathbf{a}_e$ of the cylinder projection in both cameras is found first, then projections of both cylinder end points Q_1, Q_2 are found on $\mathbf{a}_e$. The cylinder projection is searched for in the camera image inside a circular ROI (Region Of Interest), defined as in Figure 6(a): For each $\mathbf{x}_c$, a bounding sphere S_B of radius R_B equal to physical cylinder length L is projected as a bounding circle $C_B : (\mathbf{c}_B, r_B)$ to each camera plane. Note that the ROI is not necessarily centered at the center of the cylinder projection. At each circular ROI, the axis $\mathbf{a}_e$ is calculated as the first principal component of PCA (Principal Component Analysis) over all pixel positions in the ROI weighted by the pixel intensities, with $\mathbf{c}_e$ as a centroid. The circular ROI is further trimmed with an oriented bounding box $B_e : (\mathbf{c}_B, \mathbf{a}_e, l_e, w_e)$, where $l_e = 2 \cdot r_B$ and w_e is estimated conservatively (Figure 6(b)) by projecting a sphere of radius $R_C = \frac{D}{2}$ to camera plane p shifted from $\mathbf{x}_c$ towards the camera origin O by the physical cylinder length L, and D is the physical cylinder diameter. Inside the bounding box B_e the projection axis is re-estimated once again by weighted PCA. For further processing, the rectangular ROI is aligned to the cylinder projection axis, see Figure 7(a).

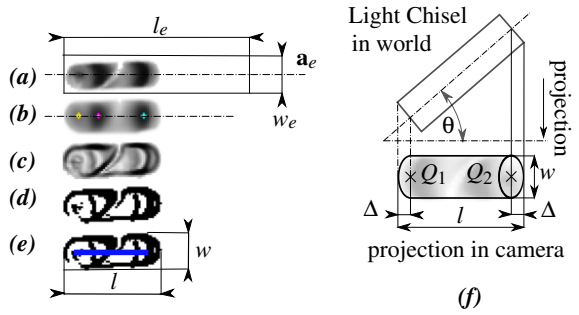


Figure 7: *Segmentation of the Light Chisel projection:* ***(a)*** *Oriented bounding box of the cylinder projection in the camera space,* ***(b)*** *Detected high intensity blobs,* ***(c)*** *Gradient magnitude,* ***(d)*** *Segmented gradient,* ***(e)*** *Length and width measured from the connected component,* ***(f)*** *Correction of the cylinder projection end points.*

The conservatively estimated B_e may contain spurious light spots, which are often outside of the Light Chisel work space. We ensure that the cameras are not overexposed, so we are able to segment the image by the hue of the cylinder emitted light. The tight cylinder projection length l and width w are measured from a connected component of the segmented gradient, seed-filled from the highest intensity blob inside B_e. Following 7 steps are performed:

1) High intensity blobs are detected along $\mathbf{a}_e$ by an LoG detector, see Figure 7(b). Note that these blobs may be different from those detected in step I. Here the LoG kernel size is tuned to the width of the cylinder projection w_e, while in step I the projection width is not know yet. 2) Similarly to [SSR08], the hue of the highest intensity blob is sampled and the pixels inside B_e with hue far from the highest intensity sample are zeroed and the blobs are re-evaluated. In this step, the pose candidates with dissimilar hues for the left / right cameras are dropped. 3) Blobs with intensity below a fraction of the highest intensity blob are dropped. One to three blobs are expected to be detected along $\mathbf{a}_e$ inside the

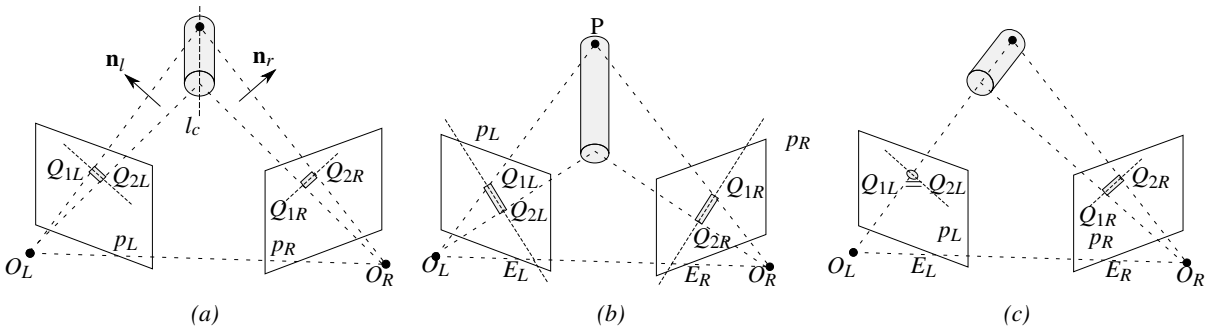


Figure 8: *The Light Chisel projects into two camera planes.* ***(a)*** *Each cylinder end point lies in a distinct epipolar plane.* ***(b)*** *Both cylinder end points project onto a common epipolar plane.* ***(c)*** *The cylinder is perpendicular to the left camera plane.*

cylinder projection, see Figure 6(c), dependent on angle θ of the cylinder axis towards the camera plane. 4) The gradient magnitude of the image inside the axis-aligned bounding box is segmented as shown in Figure 7(c,d). 5) A connected component is found touching the highest intensity blob (Figure 7(b), magenta). The black spiral wound over the diffuse cylinder may break the segmented gradient (Figure 7(d)) into two pieces, we therefore connect the maximum three blob centers surviving step 3) by a blue bar (see Figure 7(e)) before performing a seed fill over the black and blue pixels. 6) Length l of the connected component along $\mathbf{a}_e$ and width w perpendicular to $\mathbf{a}_e$ are measured. 7) Finally, cylinder projection end points Q_1, Q_2 are estimated as in Figure 7(f) from w and l, approximating the perspective projection by orthographic projection. Angle θ of the cylinder towards the camera plane is estimated by solving the quadratic equation $(v^2+1)\cos(\theta)^2 - 2uv\cos(\theta) + u - 1 = 0$ for $\cos(\theta)$, where $u = l/w$ and $v = L/D$ is the ratio of the physical cylinder length to its diameter. The correction offset Δ in the camera image from the edge of the cylinder projection towards the cylinder end point projection is calculated as $\Delta = w\sin(\theta)$.

The intensity of the cylinder pixels in the projection depends on the luminous flux distribution of the LEDs and on the cylinder material scattering characteristics. The intensity of the light captured by the camera decreases with rising angle θ of the cylinder axis towards the camera plane, and the cylinder tip is only weakly illuminated by the rear side of the front LED, negatively influencing detection of the connected component, as in Figure 7(e). Therefore, the detection quality of Q_1 and Q_2 decreases with rising angle θ.

III. 3D position / orientation (5DOF) of the Light Chisel from left / right cylinder projections: The 5DOF pose of the cylinder and its length are calculated for each cylinder position candidate from its end point projections $Q_{1L/R}, Q_{2L/R}$. Candidates with the estimated cylinder length not matching the physical length L are dropped. We found the detection of the cylinder projection axis $\mathbf{a}_e$ to be significantly more reliable than the detection of the cylinder end points $Q_{1/2}$. To utilize this a priori knowledge, we recognize three situations: 1) Two distinct cylinder end point projections are detected per camera and each 3D cylinder end point lies on a distinct epipolar plane, as in Figure 8(a). This situation of the three produces results with the best accuracy. 2) Two distinct cylinder end points are detected per camera, and the 3D cylinder axis lies in a single epipolar plane, as in Figure 8(b). 3) One end point projection is detected for one camera, and two end point projections are detected for the other camera, as in Figure 8(c). The Light Chisel is oriented nearly perpendicular to left camera.

In case 1), the 3D cylinder axis is found as the intersection of planes (O_L, Q_{1L}, Q_{2L}) and (O_R, Q_{1R}, Q_{2R}). The 3D position of the Light Chisel is calculated by projecting the cylinder end points $Q_{1L/R}, Q_{2L/R}$ onto the cylinder 3D axis from the camera, where the end point projections are calculated with higher accuracy: The camera with a smaller angle θ of the cylinder axis towards its plane is used.

In case 2), where the 3D cylinder axis lies on a single epipolar plane, the intersection of planes (O_L, Q_{1L}, Q_{2L}) and (O_R, Q_{1R}, Q_{2R}) becomes singular. Then 3D cylinder end point Q_1 resp. Q_2 is calculated as a closest point of rays (O_L, Q_{1L}) and (O_R, Q_{1R}), resp. (O_L, Q_{2L}) and (O_R, Q_{2R}). There is ambiguity in the correspondence between the left Q_{1L}, Q_{2L} and right Q_{1R}, Q_{2R} cylinder end point projections. We hoped that one of the two candidates would produce an invalid 3D cylinder length, but because of inaccuracies in the detection of the cylinder end point projections along the projection axis this was not reliable. We therefore produce two alternate 5DOF poses for this case and select the pose with the better template matching response in stage IV. Case 3) is similar to case 2), only there is no ambiguity in point correspondence, as the cylinder projects to a single blob on one camera image.

IV. Template matching: The cylinder 5DOF pose is improved and rotation of the cylinder around its axis (roll) is recovered by template matching. We render a diffuse cylinder model to the left / right camera planes and evaluate the

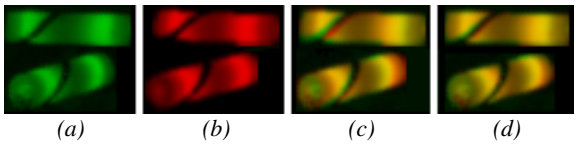


Figure 9: *Template matching against the left / right camera:* ***(a)*** *The Light Chisel captured by two cameras.* ***(b)*** *The cylinder model rendered into camera images.* ***(c)*** *The cylinder model rendered over the camera captured image, before template matching, the red and green regions show a mismatch.* ***(d)*** *After template matching, the red and green regions are reduced.*

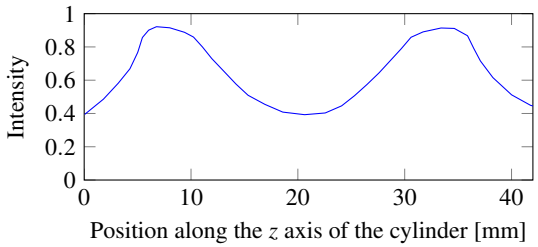


Figure 10: *Intensity profile of the emitted light along the axis of the Light Chisel cylinder when looking from the side, showing two distinctive peaks of the two LEDs.*

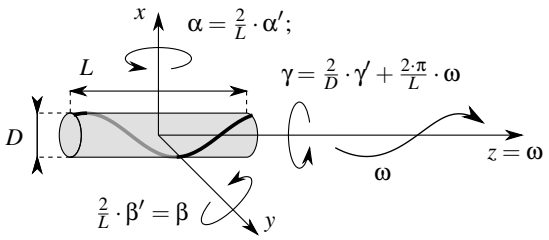


Figure 11: *Coordinate system for template matching. The cylinder pose defined by the center (x,y,z) and the Euler angles (α,β,γ) is optimized by template matching in a preconditioned space $\mathbf{x}' = (x,y,\omega,\alpha',\beta',\gamma')$, where all axes produce a similar template matching response.*

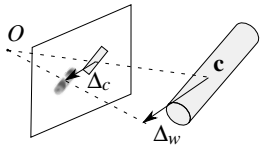


Figure 12: *Improving the cylinder detection pose by a sliding window template matching in the camera plane.*

match against the captured image. Figure 9 shows a camera captured image, a synthetic back projection to the camera image and their overlap before and after template matching.

We model the Light Chisel diffuse cylinder by mapping an emissive 1-dimensional texture along the cylinder length and onto the cylinder front side. The cylinder rear side is embedded into the tool handle and therefore does not emit light. We extracted a light intensity profile from images of the Light Chisel positioned parallel to the camera plane, captured by the tracking cameras. In these images we extracted the light intensity profile along the cylinder projection axis, see Figure 10. To improve the match, we simulate the camera off focus blur and diffraction by low pass filtering the synthetic image with a 5×5 Gaussian kernel of $\sigma = 1.1$.

6DOF pose $\mathbf{x} \in \mathbb{R}^6$, $\mathbf{x} = (x,y,z,\alpha,\beta,\gamma)$ (Figure 11) of the Light Chisel is varied by Conjugate Gradient Descent in discrete steps towards the minimum of the sum of the squared differences m between the camera image $\mathbf{I}$ and the back projected template $\mathbf{T}$, where $m = \Sigma_{x,y}(\mathbf{I}(x,y) - \mathbf{T}(x,y))^2$. Instead of varying $\mathbf{x}$ directly, minimization is performed in a preconditioned space $\mathbf{x}' \in \mathbb{R}^6$ (see Figure 11). The rotations are scaled to produce the same maximum displacement as translations. Translation along the z axis is replaced by screw ω, which separates the search along the z axis from the search along the roll axis γ. We chose the step of the gradient descent to be an order of magnitude lower than its pixel projection in the camera image to achieve subpixel precision.

Gradient descent methods require the initial position to be close to the minimum and the matching function to be locally convex and differentiable. The initial 5DOF pose estimated by step III may not be sufficiently close to the desired minimum, and the Light Chisel roll γ is not known yet. The gradient estimate in the direction of γ is zero until the spiral of the model overlaps at least partially with the spiral of the physical object. Therefore before entering the gradient descent, γ is first estimated by a full search with an angular step equal to the spiral width. This full search for γ is accelerated by summing the squares of the differences for pixels under the model spiral only, as the other model pixels do not change. Before entering the gradient descent, we also improve the initial position of the Light Chisel by sliding the synthetic image rendered into the camera plane in a window around its center of projection, as shown in Figure 12. The cylinder 3D center $\mathbf{c}$ is shifted by $\Delta_w = \Delta_c \mathbf{c}_z$, where Δ_c is a correction vector in the normalized camera plane and c_z is the distance of the cylinder center from the camera origin.

We faced the following challenges: 1) If the Light Chisel model is not accurate, template matching produces multiple local minima close to each other with comparable numerical values. Gradient descent driven by template matching snaps to these minima randomly based on the starting position, which is influenced by various noise sources. Snapping randomly to multiple local minima manifests itself as detection jitter. 2) Accurate matching to both cameras simultane-

ously is not possible in the case of imprecise camera calibration. 3) The fast moving Light Chisel is captured at different positions by unsynchronized cameras. The rolling shutter of consumer CMOS cameras produces a similar time shift effect. The first line of the sensor is exposed nearly a full camera time frame earlier than the last line. We approach the problem by first running the gradient descent for each camera independently, then we average the partial results.

4. Experimental Results

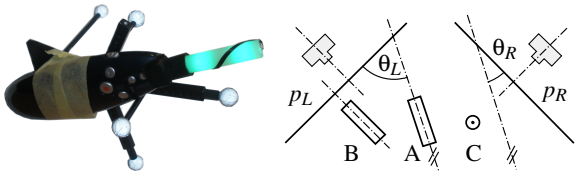
The tracking algorithm was evaluated with two consumer CMOS VGA cameras capable of 75 FPS (Sony PS3 Eye, $10 per piece), equipped with lenses of 4.5mm focal length, having the same angle of view as a lens of 41mm focal length in a 35-mm film camera. The cameras were oriented approximately perpendicularly to each other, spaced 1 m from each other and 0.7 m from the center of the work space, and their internal / external parameters were calibrated using the OpenCV library and a ring pattern. The two cameras captured a work space of roughly $560 \times 310 \times 330$ mm. The 42 mm long Light Chisel cylinder measures 46 pixels on an average when projected into the cameras, 36 pixels minimum and 73 pixels maximum. The average projection length of 46 pixels corresponds to 0.9 mm per pixel of a color image interpolated from a VGA sensor (resolution 640×480) with Bayer filter mosaic. If the emitted light activates pixels of a single color only, the effective resolution decreases to 1.8 mm per pixel for pure red and blue colors and to 1.3 mm per pixel for pure green color. Further measurements were performed with a green illuminated cylinder.

We tested our tracking setup both with and without hardware camera synchronization. Without camera synchronization, the left / right camera frames are separated up to 1/2 of a frame interval, which is equal to 6.7 ms for 75 FPS. In that interval, the Light Chisel may travel up to 5 mm in our table top setup during rapid movements, but the user has no precise control during fast movements anyway, so the resulting systematic errors may be neglected.

Our Light Chisel tracking algorithm is executed on a single thread of an Intel i7 CPU running at 3 GHz. Cylinder model rendering and template matching is accelerated by a CUDA program running on NVidia GTX470 GPU. With this setup, the tracking achieves cca. 60 updates per second. The images processed by the GPU at each step of the gradient descent are small, therefore the CUDA kernel launch cost dominates and the GPU is heavily underutilized. One could amortize the CUDA kernel launch cost by tracking multiple Light Chisels in parallel. Because the GPU is underutilized, CPU only implementation using multi-resolution methods will make the real-time CPU implementation feasible.

We evaluated our system against a set of 6 IR cameras (NaturalPoint OptiTrack S250e [Opt14] featuring a 832×832, 250 FPS sensor) tracking a rigid body with 6 retro-reflexive markers, where our Light Chisel was affixed rigidly to the tracked 6-marker body, as shown in Figure 13(a). The combo was tracked by both systems simultaneously during a typical set of 6DOF movements under typical indoor lighting conditions. The tracking sequence, over which we evaluated our tracking algorithm, is provided at the end of our video. To suppress the effects of rapid movements on absolute accuracy, we synchronized the left / right cameras of our system.

In order to compare the OptiTrack data set with ours, rigid transformation from the OptiTrack coordinate system to ours (6DOF), and also rigid transformation from the 6-marker body to the Light Chisel device coordinate system (another 6DOF) and time offset have to be known with high accuracy. We find the optimal rigid transformation from the OptiTrack coordinate system to ours and the time offset of the two trackers by solving $\operatorname{argmin}_{R,T,t} e(R,T,t)$ by gradient descent, where $e(R,T,t) = \sum_i \sum_{j=1}^{2} \left| \mathbf{P}^i_j - \bar{\mathbf{P}}_j \right|^2$, $\mathbf{P}^i_j$ is the position of the j-th Light Chisel cylinder end point inside the OptiTrack 6-ball body coordinate system in the i-th time frame, and $\bar{\mathbf{P}}_j$ is the average position of the j-th cylinder end point over all time frames. The transformation from the OptiTrack 6-ball body to the Light Chisel coordinate system is then calculated as a mean over the tracking sequence using the optimized (R,T,t) parameters. The tracking data captured by OptiTrack appear smooth, but we experienced some sudden jumps in the range of 1 to 2 mm as groups of retro-reflexive balls move from the range of one set of OptiTrack cameras to the other.



(a) (b)

Figure 13: ***(a)*** *The Light Chisel affixed to an OptiTrack target with 6 retro-reflexive balls.* ***(b)*** *Classification of Light Chisel angles $\theta_{L,R}$ towards the camera planes $p_{L,R}$. Cylinder A is in a general pose. Cylinder B is placed perpendicular to p_L ($\theta_L = 90^\circ$). C is placed parallel to both cameras ($\theta_L = \theta_R = 0^\circ$).*

Figure 14 shows the absolute position and orientation of a Light Chisel in world coordinates over a sample interval of 3 seconds, detected by our system. Figure 15 visualizes the difference in the position and orientation of the same sequence referenced to OptiTrack. All difference curves show a nonzero mean over the 3-second interval, but the OptiTrack trajectory was aligned to the LightChisel trajectory for zero mean over a 35-second tracking interval. The short term nonzero mean may indicate a systematic error of our

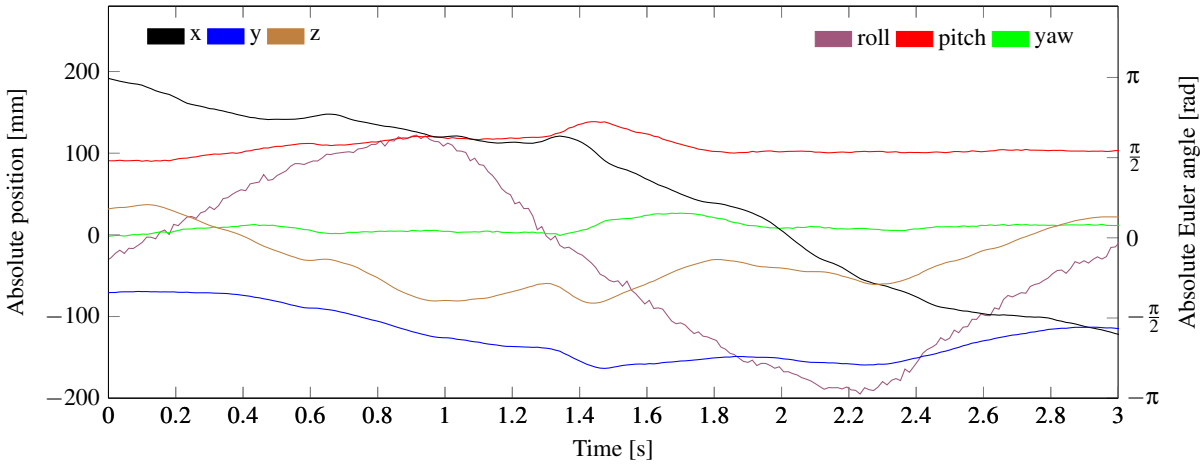


Figure 14: *Absolute position and orientation of the Light Chisel over a sample interval of 3 seconds, detected by our system*

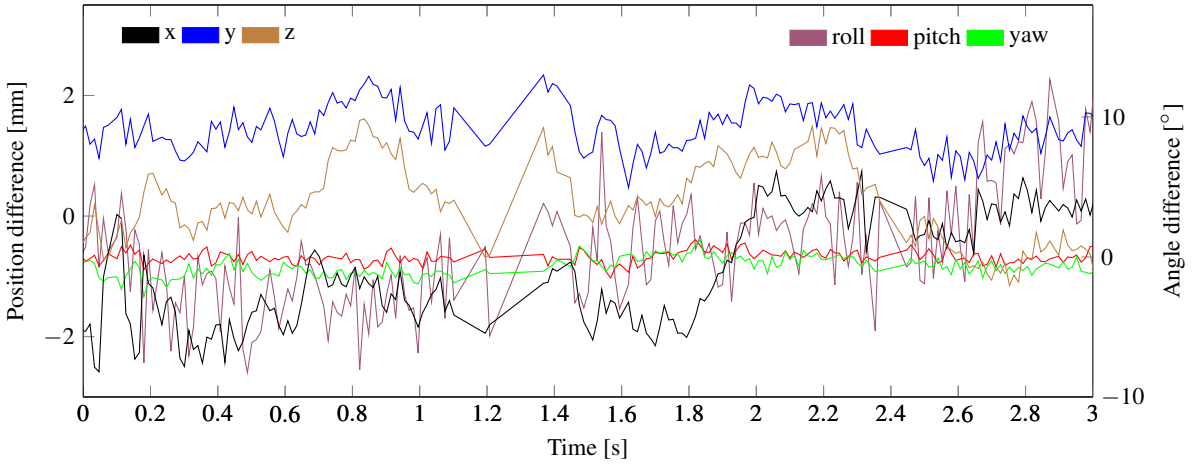


Figure 15: *Difference of position and orientation from the OptiTrack reference over a 3-second time interval. In the time intervals (1.1s, 1.4s) and (2.35s, 2.45s), OptiTrack reported low quality of tracking, so the OptiTrack data were ignored.*

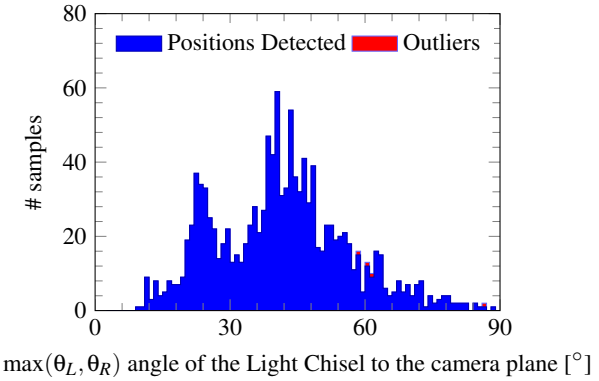


Figure 16: *Histogram of # of the Light Chisel position samples per maximum angle to the camera plane*

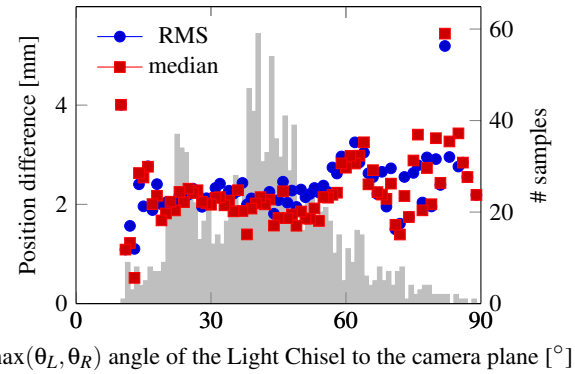


Figure 17: *Difference of the Light Chisel position detected by our system from the OptiTrack reference*

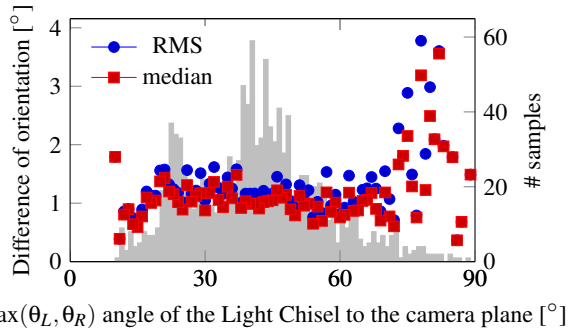


Figure 18: *Difference of the Light Chisel orientation (yaw + pitch) detected by our system from the OptiTrack reference*

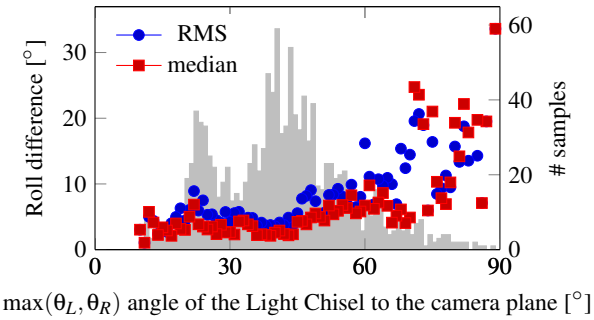


Figure 19: *Difference of the Light Chisel roll detected by our system from the OptiTrack reference*

system as well as of Optitrack. The roll curve shows higher jitter, which we attribute to the effect of the Light Chisel spiral on the template matching response. Starting from a perfect alignment pose and screwing the cylinder around the ω axis, as shown in Figure 11, produces a template matching error mainly at the ends of the cylinder, while not changing the template matching response along the spiral projection. Screwing the cylinder from a perfect alignment pose by Δ_z ($= \Delta_\omega$) produces an error in the roll pose of $\Delta_\gamma = \Delta_z \cdot \frac{2 \cdot \pi}{L}$, which corresponds to an 8.57° roll error per mm z position error. In our modeling application, we applied temporal filtering to the roll axis. The added latency was found acceptable, because the rolling movement requires fine coordination of finger movements and is therefore slow. The tracking noise in the other five axes was found to be acceptable without temporal filtering.

The Light Chisel detection success is shown in Figure 16 based on the maximum angle $max(\theta_L, \theta_R)$ (see Figure 13(b)) of the cylinder axis towards the camera planes. A smaller angle θ results in longer projection in the camera space and therefore better detection. A Light Chisel pose is classified as an outlier if the difference against the OptiTrack reference is higher than 40 mm in position, 30° in yaw or pitch and higher than 45° in roll. Out of the 1239 tracking samples, 4 were classified as outliers. The outlier poses are usually identifiable as sudden jumps in the 6DOF trajectory and may be interpolated from the neighboring positions. RMS, mean and average differences from the OptiTrack reference over the 1239 tracking samples are listed in following table:

	RMS	mean	median
position [mm]	2.25	2.07	2.02
yaw and pitch [°]	1.26	1.06	0.93
roll [°]	7.25	5.26	3.82

Figure 17 shows the absolute difference of the Light Chisel position in relation to the position calculated by OptiTrack. A position difference greater than the camera pixel resolution may be explained by an imprecise camera calibration and distortion model, by an imprecise Light Chisel cylinder model, and by OptiTrack tracking errors. Figure 18 shows deviation of the Light Chisel orientation vector from the orientation detected by OptiTrack, Figure 19 depicts the deviation of the Light Chisel roll from the roll detected by OptiTrack.

The position, orientation and roll detection errors increase with cylinder axis angle θ towards the camera plane increasing above 70° (shown in Figure 13(b), pose B), corresponding to 6% area on a unit hemisphere. This is to be expected, as the foreshortening of the cylinder projection in one camera makes detection of the cylinder projection axis less reliable. We also found out, that the light intensity profile, as shown in Figure 10 for $\theta = 0°$, changes to some extent with the angle θ. Using a light intensity profile measured at $\theta = 0°$ makes template matching for $\theta \approx 90°$ less reliable.

4.1. Robustness

We tested our Light Chisel indoors under similar ambient light conditions as in IrPen [HHLL14]. In the Applications section of our video, the room is kept dark for filming only, to improve contrast of the image projected on the half transparent mirror of our AR setup. The last section of our video shows the performance of our tracking algorithm under normal indoor lighting conditions, under which the evaluation against OptiTrack was performed. The right side of the left camera view is cluttered with a white check board lit by a direct sun light.

The intensity of the light emitted by the Light Chisel was set to be significantly higher than the ambient light, yet still comfortable to the eyes. The use of the Light Chisel outdoors requires the ratio of Light Chisel intensity to the ambient

light to be maintained, as in the case of other optical systems. We verified, that a single setting of the Light Chisel intensity and camera exposition parameters is adequate for a range of indoor conditions from complete darkness to fluorescent light or indirect sunlight with some speckles of direct sun light in the camera view.

We also verified, that our system effectively rejects spurious objects due to filtering by the amplitude of the LoG blob detector, by epipolar constraints, by matching the hue of the left / right blob, by hue segmentation and by physical cylinder length constraints. Most of the spurious objects or light reflections are placed outside of the work space, so they are seen by a single camera only, where the probability of finding a match in the other camera is low.

4.2. Applications

We tested the Light Chisel in a seated see-through augmented reality setup similar to SpaceTop [LOIB13] and Spinnstube [WRG07]. We used the see-through setup in application for virtual sculpting (similar to [GH91]) and painting, where the Light Chisel is used similarly to a real chisel or brush. Our Light Chisel is particularly interesting for modeling swirls [ACWK04], allowing the user to control the pulling of the material and the position, orientation and velocity of a swirl simultaneously. Apart from the IR-Pen [HHLL14], which we find difficult to reproduce, we know of no other tool, that would enable modeling swirls with the easiness of our Light Chisel. We also applied our Light Chisel for controlling handle based biharmonic deformation [BK04] using libigl [JP*13]. A simple character animation is achieved by marking / unmarking the handle regions and by manipulating the handle regions in 6DOF by our Light Chisel. Suitability to perform 6DOF rigid transformations was demonstrated on an assembly task of a differential mechanism.

Manipulating a 6DOF tool in 3D space provides a much better metaphor to the applications we evaluated than the keyboard + mouse or a 2D touch based UI. The usual mouse driven UI for 3D input requires a lengthy training. The user has to learn how to manipulate components using a virtual track ball, how the mouse cursor projects on a selected 3D plane or a 3D object, how to manipulate the camera, and how to switch between these modes of operation. The keyboard + mouse does not really map well to the 3D world. On the contrary, our tool is adopted intuitively by inexperienced users in our example applications after a brief demonstration. In AR, the projection of the virtual world is aligned with the physical tool, so the visual perception is aligned with the perception of the hand motion, increasing the acceptance.

4.3. Limitations

The experimental results show decreased accuracy and reliability of tracking when the cylinder axis approaches a perpendicular orientation towards one of the camera planes. Tracking may be improved significantly by adding a third camera oriented perpendicular to the other two. In the general case, the Light Chisel pose will be calculated from the pair of cameras with the longest cylinder projection. In the case shown in Figure 8(b), where the cylinder axis lies in an epipolar plane of a camera pair, the camera of that pair with the shorter cylinder projection is disregarded.

Our algorithm requires manual exposure control of the cameras to avoid over-exposition by the Light Chisel. As an alternative to the Sony PS3 Eye camera, we considered cheap web cams. From three randomly picked low end web cams, only the Logitech C170 offered manual exposure control.

5. Conclusion and Future Work

We have presented a low-cost lightweight light pen device and a method of its 6DOF tracking by two commodity cameras. A prototype of our Light Chisel device can be built for less than $10 in material costs. Our system achieved mean differences from OptiTrack reference of 2.07 mm in position, 1.06° in yaw and pitch, and 5.26° in roll using a pair of VGA cameras. We have compared our method to other systems capable of 6DOF tool tracking, and we have shown our low cost solution to be practical for tasks demanding 6DOF, e.g. geometric modeling or character animation. Our tool may be used directly in systems equipped with a pair of cameras, e.g. the PS4 console, or with a laptop with a built-in front-facing camera and one additional external camera.

In future work, we will improve the camera calibration with the work of Datta and Jun-Sik [DKK09], and we plan to test the continuous update of the camera extrinsic matrices as in [WB97], to counteract small but gradual changes in camera positions due to the component aging and physical abuse. We also plan to improve the accuracy for unsynchronized cameras during rapid movements, using Kalman filters for prediction.

Acknowledgements

We thank Josef Havran, Jan Hošek, Miroslav Roubal and Miroslav Skrbek for their effort on the hardware of the Light Chisel and of our AR setup. We thank Jiří Bittner, Robin Healey, Ondřej Jamriška and Daniel Sýkora for proofreading. We are indebted to David Sedláček for help with OptiTrack and with the video. We thank all the anonymous reviewers for their insightful comments, which helped to improve our paper. We thank the Stanford 3D Scanning Repository for the model of the Armadillo man. Dragos V. Mitrofan allowed us to use his model of differential mechanism in our video. Our work was funded by the Czech Science Foundation under the research project P202/12/2413 (Opalis) and it builds on the results of the EU-IST project number IST027039 (ARISE) funded by the European Union.

References

[ACWK04] ANGELIDI A., CANI M.-P., WYVILL G., KING S.: Swirling-sweepers: constant-volume modeling. In *Computer Graphics and Applications, 2004. PG 2004. Proceedings. 12th Pacific Conference on* (Oct 2004), pp. 10–15. 11

[BK04] BOTSCH M., KOBBELT L.: An Intuitive Framework for Real-time Freeform Modeling. In *ACM SIGGRAPH 2004 Papers* (New York, NY, 2004), SIGGRAPH '04, ACM, pp. 630–634. 11

[DKK09] DATTA A., KIM J.-S., KANADE T.: Accurate Camera Calibration using Iterative Refinement of Control Points. In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on* (Sept 2009), pp. 1201–1208. 11

[GH91] GALYEAN T. A., HUGHES J. F.: Sculpting: An Interactive Volumetric Modeling Technique. *SIGGRAPH Comput. Graph. 25*, 4 (July 1991), 267–274. 11

[GJP*14] GUNA J., JAKUS G., POGAČNIK M., TOMAŽIČ S., SODNIK J.: An Analysis of the Precision and Reliability of the Leap Motion Sensor and Its Suitability for Static and Dynamic Tracking. *SENSORS 14*, 2 (2014), 3702–3720. 3

[HB01] HIGHTOWER J., BORRIELLO G.: Location Systems for Ubiquitous Computing. *Computer 34*, 8 (Aug. 2001), 57–66. 2

[HHC*11] HEO S., HAN J., CHOI S., LEE S., LEE G., LEE H.-E., KIM S., BANG W.-C., KIM D., KIM C.: IrCube Tracker: An Optical 6-DOF Tracker based on LED Directivity. In *Proc. of the 24th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, 2011), UIST '11, ACM, pp. 577–586. 3

[HHLL14] HAN J., HEO S., LEE H.-E., LEE G.: The IrPen: A 6-DOF Pen for Interaction with Tablet Computers. *IEEE Computer Graphics and Applications 34*, 3 (2014), 22–29. 3, 10, 11

[Hol13] HOLZ D.: Motion Capture using Cross-Sections of an Object. US Patent Appl. US 2013/0182079 A1, July 2013. 3

[HZ03] HARTLEY R., ZISSERMAN A.: *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, New York, NY, 2003. 4

[JP*13] JACOBSON A., PANOZZO D., ET AL.: libigl: A simple C++ geometry processing library, 2013. http://igl.ethz.ch/projects/libigl/. 11

[JPG*14] JACOBSON A., PANOZZO D., GLAUSER O., PRADALIER C., HILLIGES O., SORKINE-HORNUNG O.: Tangible and Modular Input Device for Character Articulation. *ACM Trans. Graph. 33*, 4 (July 2014), 82:1–82:12. 3

[LOIB13] LEE J., OLWAL A., ISHII H., BOULANGER C.: SpaceTop: Integrating 2D and Spatial 3D Interactions in a See-through Desktop Environment. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, 2013), CHI '13, ACM, pp. 189–192. 11

[MG03] MATVEYEV S. V., GÖBEL M.: The Optical Tweezers: Multiple-Point Interaction Technique. In *Proc. of the ACM Symposium on Virtual Reality Software and Technology* (New York, NY, 2003), VRST '03, ACM, pp. 184–187. 3

[MIF12] MCILROY P., IZADI S., FITZGIBBON A.: Kinectrack: Agile 6-DoF Tracking Using a Projected Dot Pattern. In *Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on* (Nov 2012), pp. 23–29. 3

[MWH*09] MOHAN A., WOO G., HIURA S., SMITHWICK Q., RASKAR R.: Bokode: Imperceptible Visual Tags for Camera Based Interaction from a Distance. In *ACM SIGGRAPH 2009 Papers* (New York, NY, 2009), SIGGRAPH '09, ACM, pp. 98:1–98:8. 3

[Opt14] OPTITRACK: OptiTrack S250e, 2014. http://www.naturalpoint.com/optitrack/products/s250e/. 2, 8

[PK08] PINTARIC T., KAUFMANN H.: A Rigid-Body Target Design Methodology for Optical Pose-Tracking Systems. In *Proc. of the 2008 ACM Symposium on Virtual Reality Software and Technology* (New York, 2008), VRST'08, ACM, pp. 73–76. 2

[QSW*14] QIAN C., SUN X., WEI Y., TANG X., SUN J.: Realtime and Robust Hand Tracking from Depth. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on* (June 2014), pp. 1106–1113. 3

[RNd*07] RASKAR R., NII H., DEDECKER B., HASHIMOTO Y., SUMMET J., MOORE D., ZHAO Y., WESTHUES J., DIETZ P., BARNWELL J., NAYAR S., INAMI M., BEKAERT P., NOLAND M., BRANZOI V., BRUNS E.: Prakash: Lighting Aware Motion Capture using Photosensing Markers and Multiplexed Illuminators. *ACM Trans. Graphics 26*, 3 (July 2007), 36:1–36:11. 3

[SSR08] SÝKORA D., SEDLÁČEK D., RIEGE K.: Real-time Color Ball Tracking for Augmented Reality. In *Proc. of Eurographics Symposium on Virtual Environments* (2008), pp. 9–16. 5

[TSLP14] TOMPSON J., STEIN M., LECUN Y., PERLIN K.: Real-Time Continuous Pose Recovery of Human Hands Using Convolutional Networks. *ACM Trans. Graph. 33*, 5 (Sept. 2014), 169:1–169:10. 3

[WB97] WELCH G., BISHOP G.: SCAAT: Incremental Tracking with Incomplete Information. In *Proc. of SIGGRAPH 97* (Aug. 1997), pp. 333–344. 3, 11

[WBRF13] WEICHERT F., BACHMANN D., RUDAK B., FISSELER D.: Analysis of the Accuracy and Robustness of the Leap Motion Controller. *SENSORS 13*, 5 (May 2013), 6380–6393. 3

[WF02] WELCH G., FOXLIN E.: Motion Tracking: No Silver Bullet, but a Respectable Arsenal. *Computer Graphics and Applications, IEEE 22*, 6 (Nov 2002), 24–38. 2

[WNG*06] WIENSS C., NIKITIN I., GOEBBELS G., TROCHE K., GÖBEL M., NIKITINA L., MÜLLER S.: Sceptre: An Infrared Laser Tracking System for Virtual Environments. In *Proc. of the ACM Symposium on Virtual Reality Software and Technology* (New York, NY, 2006), VRST '06, ACM, pp. 45–50. 3

[Wor13] WORLDWIZ: Interaction Wand for Virtual Reality, 2013. http://www.worldviz.com/products/ppt/wand-and-eyes. 3

[WP09] WANG R. Y., POPOVIĆ J.: Real-Time Hand-Tracking with a Color Glove. *ACM Trans. on Graphics 28*, 3 (July 2009), 63:1–63:8. 3

[WPP11] WANG R., PARIS S., POPOVIĆ J.: 6D Hands: Markerless Hand-tracking for Computer Aided Design. In *Proc. of the 24th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, 2011), UIST '11, ACM, pp. 549–558. 3

[WTK13] WANG R. Y., TWIGG C., KIN K.: 3Gear Systems, 2013. http://www.threegear.com/. 3

[WRG07] WIND J., RIEGE K.,BOGEN M.: Spinnstube®: A Seated Augmented Reality Display System. In *Proc. of the 13th Eurographics conference on Virtual Environments* (Weimar, Germany, 2007), EGVE '07, Eurographics, pp. 17–23. 11

[XHW*13] XIAO R., HARRISON C., WILLIS K. D., POUPYREV I., HUDSON S. E.: Lumitrack: Low Cost, High Precision, High Speed Tracking with Projected m-Sequences. In *Proc. of the 26th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, 2013), UIST '13, ACM, pp. 3–12. 3

[zSp13] ZSPACE: zSpace System, 2013. http://www.zspace.com/. 3