Super-Resolution Capacitive Touchscreens

Sven Mayer
Carnegie Mellon University
Pittsburgh, PA, USA
info@sven-mayer.com

Xiangyu Xu
Carnegie Mellon University
Pittsburgh, PA, USA
xuxiangyu2014@gmail.com

Chris Harrison
Carnegie Mellon University
Pittsburgh, PA, USA
chris.harrison@cs.cmu.edu

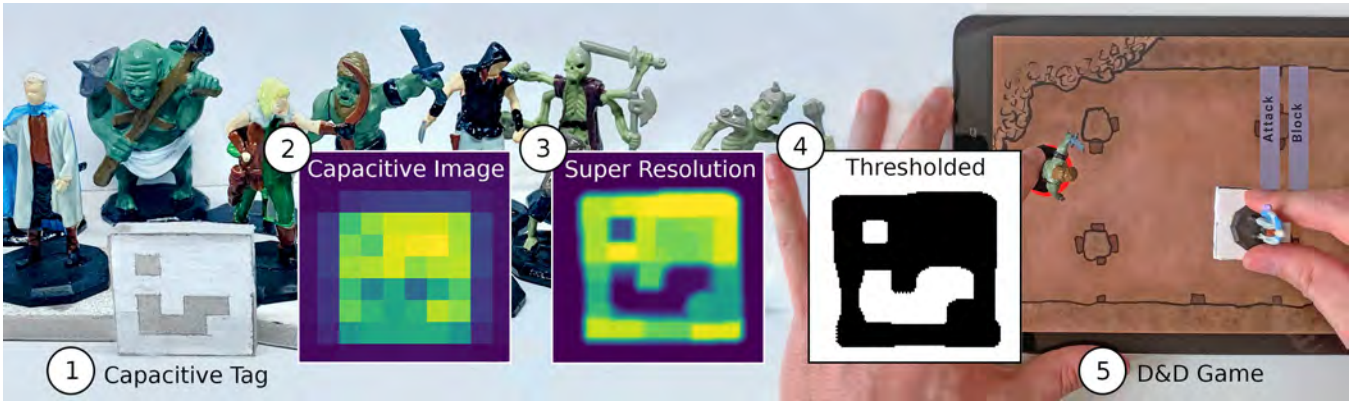

Figure 1: Today's capacitive touchscreen are low resolution, precluding recognition of e.g., passive tangibles (1). By capturing several raw frames (2) and applying super-resolution techniques, resolution can be dramatically improved (3,4), enabling new interactive experiences (5).

ABSTRACT

Capacitive touchscreens are near-ubiquitous in today's touch-driven devices, such as smartphones and tablets. By using rows and columns of electrodes, specialized touch controllers are able to capture a 2D image of capacitance at the surface of a screen. For over a decade, capacitive "pixels" have been around 4 millimeters in size – a surprisingly low resolution that precludes a wide range of interesting applications. In this paper, we show how super-resolution techniques, long used in fields such as biology and astronomy, can be applied to capacitive touchscreen data. By integrating data from many frames, our software-only process is able to resolve geometric details finer than the original sensor resolution. This opens the door to passive tangibles with higher-density fiducials and also recognition of every-day metal objects, such as keys and coins. We built several applications to illustrate the potential of our approach and report the findings of a multipart evaluation.

CCS CONCEPTS

• **Human-centered computing** → **Touch screens**.

CHI '21, May 8–13, 2021, Yokohama, Japan

ACM ISBN 978-1-4503-8096-6/21/05…$15.00
https://doi.org/10.1145/3411764.3445703

KEYWORDS

Capacitive sensing, Super-resolution, Tangibles, Tracking, Touch input

ACM Reference Format:
Sven Mayer, Xiangyu Xu, and Chris Harrison. 2021. Super-Resolution Capacitive Touchscreens. In *CHI Conference on Human Factors in Computing Systems (CHI '21), May 8–13, 2021, Yokohama, Japan.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3411764.3445703

1 INTRODUCTION

Touch input is the dominant mode of interaction on mobile computers and is increasingly common in laptops, self-service kiosks, domestic appliances, and cars. The most prevalent touchscreen technology used today is projective-capacitive [14, 53], which uses a row-column arrangement of sensing electrodes (typically made from transparent conductive materials, such as ITO, sandwiched between cover glass and a display). A specialized sensing IC measures the coupling capacitance not only at a single point (e.g., [38, 68]), but at each row-column intersection [9, 42], building a 2D signal often called a "capacitive image" in the literature [22, 30, 60]. Capacitive objects (e.g., fingers and metal items) touching the screen appear as "blobs" in the image, which can be tracked over time by standard computer vision algorithms, enabling inputs such as taps and swipes.

As touchscreens are primarily designed to capture finger input, the pitch of the capacitive matrix is generally sized such that fingertips overlap at least two pixels horizontally and vertically, as this permits a fairly accurate sub-capacitive-pixel interpolation of the true touch centroid. For this reason, capacitive touchscreen sensor resolution has not changed much over time or device category:

e.g., LG 'G' smartwatch (3.5mm pitch) [60], Nexus 5 smartphone (4.1mm pitch) [31], Samsung S4 smartphone (3.9mm pitch) [60], Samsung Galaxy Tab S2 tablet (4.0mm pitch), and Microsoft 55″ Surface touchscreen (~5.9mm pitch) [59]. This coarse resolution immediately precludes many interesting applications. Moreover, increasing capacitive touchscreen sensor resolution is not trivial, as adding rows and columns incurs a quadratic cost in terms of the number of intersections to sense, increasing latency when the current trend is towards more responsive touchscreens. Thus, if we wish to have higher-resolution capacitive image data, we must turn to other approaches.

In this paper, we show how super-resolution techniques – long used in fields such as biology and astronomy – can be applied to capacitive images. The fundamental operation begins by capturing a series of images at slightly different perspectives or offsets. In the case of imaging a celestial body, this might be different perspectives as the Earth orbits the Sun. In our case, it is translations of an object on a touchscreen's surface that quantizes the object along many different capacitive pixel boundaries. Although single frames inherently contain no details smaller than a pixel, sub-pixel details can now be resolved when fused together through super-resolution. Fortunately, there is often sufficient sub-pixel movement created when a user naturally places an object down onto a touch surface. However, this does come at the cost of latency, as more than one frame is needed. For this reason, we envision a single-frame touch-tracking pipeline and multi-frame super-resolution pipeline running in parallel, maintaining existing touch input responsiveness while also opening new interactive opportunities.

Figure 1 offers an example result: a $4{\times}4mm$ AprilTag fiducial [54] made from metal foil is unrecoverable with the native capacitive image. However, with just a few frames captured during translation of the tangible (perhaps when first placed down onto the tablet, or when the figurine is moved), super-resolution can reveal the tag's fine details, permitting recognition with a standard AprilTag reader. Similar results are achieved for human inputs, such as palm prints (see Figure 2) and everyday conductive objects, such as coins and keys; see Figure 4. Importantly, our software could be made to run on nearly all smartphones, unlocking new interactive capabilities without any changes to the hardware.

2 RELATED WORK

Our work draws on several disparate literatures, most notably super-resolution techniques, capacitive touchscreen, and tangibles, which we now review.

2.1 Super-Resolution Techniques

Super-resolution techniques have been developed and applied in many domains, including traditional photography [58], astronomy [23, 41], CCTV footage [56], medical imaging [15], and microscopy [19]. The universal goal across all these domains is to enhance resolution of images beyond that of the native sensor. Traditional geometric approaches, such as [10, 11, 17], rely on aligning and stacking multiple frames to mitigate sensor noise and resolve finer-grained details. Other super-resolution techniques use time varying information, such as the transient fluorescence of molecules under an optical microscope [2, 19].

By applying domain knowledge (e.g., the approximate 3D geometry of human faces [56], approximate finger shape for superior touch centroid prediction [64]), super-resolution output can be further improved. There are also many deep-learning-based super-resolution approaches, which learn to infill fine-grained geometric detail from a training corpus of domain-specific images [8, 9, 63, 65, 68]. Such a deep learning approach is not yet feasible for touchscreen capacitive images as there are no large corpora for training, nor do we have sufficient starting resolution to perform the downsample-upsample-and-compare-style training procedure most often employed in these papers. The latter uses multi-megapixel photos, while the capacitive images on our test tablet is a mere 49×37 pixels (<0.002 of a megapixel). For this reason, we utilize a purely geometric approach. Finally, super-resolution imaging has fundamental limits, and there is considerable theoretical research in this space [1, 35].

2.2 Touchscreen "Capacitive Images"

Grosse-Puppendahl et al. [16] provides an excellent review of uses of capacitive sensing in Human-Computer interaction (HCI). More relevant to the present work are efforts that specifically leveraged the capacitive image. For example, Kumar et al. [30] used the data to improve touch precision over the algorithms run on the touch controller. The capacitive image has also been used to enable new interactive modalities, including estimating a finger's 3D orientation on a screen [36, 60], and even recognizing digits (e.g., thumb, pinky) [33]. Holz et al. [22] used the capacitive image to recognize users' ears, and later, Guo et al. [18] showed handprints could also be used to differentiate users in small groups.

While we did not reimplement any of these systems, we believe our super-resolution approach would almost certainly improve the accuracy of these prior systems. For instance, Guo et al. [18] used raw capacitive images of palm prints for biometrics (example "raw image" shown in Figure 2), which are greatly improved with super resolution (Figure 2, "super-resolution output"); Not only are characteristic contours revealed, but more accurate measurements of e.g., finger lengths are also possible. Note that interpolation/smoothing is commonly applied to raw capacitive images, but does not reveal any new information or the same level of geometric details (Figure 2 offer interpolated examples for reference).

2.3 Tangibles

Although tangibles (physical objects that can be manipulated by users on a screen and participate in the interactive experience) are not the focus of our innovation, they are nonetheless enabled by our approach in a higher-density manner than previously demonstrated on capacitive touchscreens. For this reason, it is worth briefly reviewing tangibles in general and prior capacitive tangible systems more specifically.

Many of the very earliest explorations of touch screens and "surface computing" incorporated tangibles; e.g., [6, 13, 25, 39, 48, 49]. In the 2000s, a wide range of "multi-touch" research systems were developed [7, 20], and in 2007, Microsoft launched a commercial "Surface" computer (later rebranded PixelSense) [37]. Most of these systems used cameras operating behind a defuse screen, able to detect user touches, and in many cases recognize tangibles, either through visual markers (e.g., [27]) or object contours (e.g., [57]).

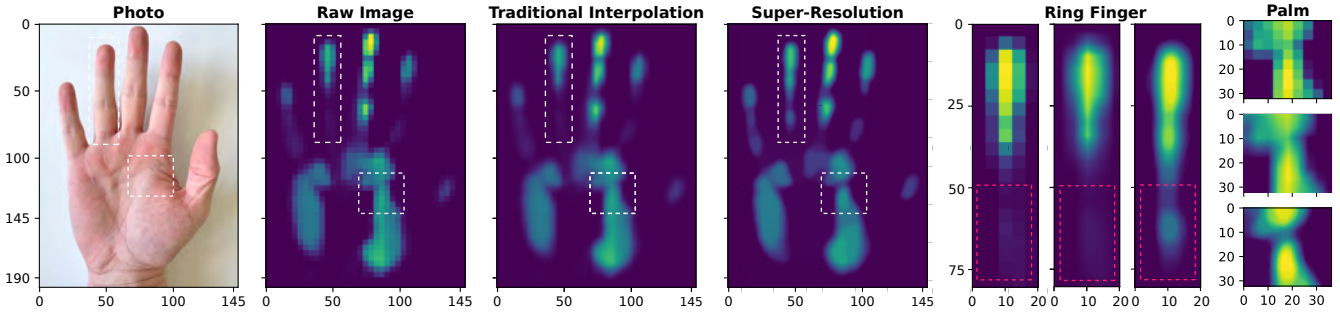

Figure 2: The left-most image is a reference photo of a user's handprint. Next are the three versions of the handprint as captured by a touchscreen: a raw capacitive image, an interpolated capacitive image, and a super-resolved capacitive image. Two areas are highlighted with white dashes; zoomed-in versions are shown far right (order is raw, interpolated, and super-resolved. Ring Finger is ordered left to right and Palm is top to bottom).

However, by the end of the decade, camera-based touch tables were largely displaced by capacitive sensing technologies, offering thinner form factors. With only a low-resolution capacitive matrix for sensing, libraries of tangibles became challenging, if not impossible to support. Researchers have explored several ways to (re)enable tangibles on standard capacitive touchscreens, including active battery-powered tags [40, 51, 52], low-density passive capacitive fiducials [5, 29, 34, 44, 51, 52], and contour recognition [50, 59]. Our super-resolution approach could immediately be used to improve the latter two categories.

3 IMPLEMENTATION

As a proof-of-concept device on which we implemented our capacitive touchscreen super-resolution pipeline, we selected a Samsung Galaxy Tab S2 tablet. The 9.1″ touchscreen on this device captures a 49×37 capacitive image (6.33PPI or 4.02mm pixels). As noted in our introduction, this capacitive pixel size is roughly average among devices and serves as a good exemplar. Commercial devices rarely expose the capacitive image via a public API, and so we deployed a custom kernel driver to directly communicate with the Synaptics touch controller (cf. Le et al. [32]). Transmitting over I^2C, our software receives capacitive images at ~16.2FPS. We note that our implementation is not device, frame rate, or resolution dependent. Indeed, pretty much every capacitive touchscreen device made in the last decade could run our software with the appropriate driver to interface with the touchscreen (or our pipeline could run on the touch controller itself).

First, we identify all blobs by finding the contours [46] on a thresholded version of the capacitive image. This threshold was calibrated per input category (e.g., keys, coins, fiducials) to produce the best output. We track blobs over time using a Euclidean-distance centroid matching algorithm (see e.g., [28, 43]) tracking methods e.g. [3, 26, 66] or even deep-learning tracking approaches, e.g., [21]. Therefore, each tracked blob is an ever-enlarging sequence of N low-resolution frames $\{Y_i\}_{i=1}^{N}$. Our algorithm, heavily based on Xu et al. [62], takes these sequences of low-resolution frames as input, which are derived from a latent high-resolution image X of the tracked object. The generation process of $\{Y_i\}_{i=1}^{N}$ can be modeled as:

$$Y_i = g(f_{A_i}(X)) + n, \tag{1}$$

where the latent image X is first transformed by a motion model f_{A_i} and then downsampled by the sampling function g (i.e., the low-resolution touchscreen). The motion model f_{A_i} is defined by a Euclidean matrix A_i accounting for both translation and rotation. n represents gaussian noise.

With the above formulation, we can solve the multi-frame super-resolution problem in a maximum-a-posteriori (MAP) framework [27, 62], which leads to the following optimization problem:

$$\min_X \sum_{i=1}^{N} \left\| Y_i - g(f_{A_i}(X)) \right\|^2 + \lambda p(X) \tag{2}$$

where the first term is derived from the likelihood of the low-resolution observations and ensures the data fidelity of the estimation. $p(X)$ is the prior term describing the sparsity of the output image. Similar to Xu et al. [62], we use the ℓ_0 gradient prior as $p(X)$ to encourage clear shape and sharp edges in the generated results. Due to the high-complexity and non-convexity of Equation (2), we split it into two simpler sub-problems to approximate the original problem:

$$\min_{X'} \sum_{i=1}^{N} \left\| Y_i - g(f_{A_i}(X')) \right\|^2, \tag{3}$$

and

$$\min_X \sum_{i=1}^{N} \left\| X - X' \right\|^2 + \lambda p(X). \tag{4}$$

The first sub-problem Equation (3) focuses on the data fidelity term, which leads to an intermediate result X' that is compatible with the observations $\{Y_i\}_{i=1}^{N}$. The second sub-problem Equation (4) ensures the final solution not only being close to X', but also possessing the sparsity modeled by $p(X)$.

To solve Equation (3), we use bicubic upsampling [24] as the inverse-downsampling operation g^{-1}. We select a seed frame Y_s by fitting a bounding rectangle [47] to all images $\{Y_i\}_{i=1}^{N}$, and Y_s is the image with the smallest area. We then use the ECC algorithm [10] to estimate each transformation A_i between the upsampled images $g^{-1}(Y_s)$ and $g^{-1}(Y_i)$. The ECC algorithm [10] can account for both

translation and rotation, while other geometric approaches such as [11, 17] only consider translation and are not suitable for objects that rotate on the screen. With g^{-1} and the estimated $\{A_i\}_{i=1}^{N}$, the solution for Equation (3) can be given by:

$$\frac{1}{N}\sum_{i=1}^{N} f_{A_i}^{-1}(g^{-1}(Y_i)), \tag{5}$$

where we essentially take the mean of the per-frame estimates to aggregate multi-frame information. For the second sub-problem Equation (4), since it has the same form as the ℓ_0 filter [61, 62], we solve it similarly with the half-quadratic splitting scheme [55].

Finally, after receiving the super-resolved capacitive image, we can more faithfully estimate the shape [45] and size [12, 47] of objects, and use any contour matching algorithms to recognize objects [4], such as keys, hands, and fingers. For better human visual perception, a deblur [14] can be applied to the super-resolved capacitive image.

4 EXAMPLE USES

Capacitive touchscreens augmented with super resolution could enhance both existing interactions and unlock entirely new experiences. To illustrate the breadth of potential uses, we developed pairs of demonstration applications across three input categories: tangibles, everyday objects, and users. We now briefly describe these apps; please also see Video Figure.

4.1 Tangibles

As discussed in Related Work, tangibles – which were previously common on camera-based touch surfaces – have almost disappeared due to lack of support on today's capacitive touchscreens. With super resolution, we show that large libraries of low-cost passive tangibles are possible again, on standard capacitive touchscreens. Specifically, we used metal foil, which could be inexpensively stamped onto solid or even disposable paper tangibles.

As a proof-of-concept, we built two games. The first is a tabletop fantasy game, with plastic hero figurines and enemy tangibles placed into a virtual dungeon, see Figure 1. For recognition, we used 16h5 AprilTags [54] $36 \times 36mm$ in size (Figure 3A). In order to receive enough unique frames to achieve super-resolution, character tangibles are translated on the screen's surface as they explore and fight. For our second example app utilizing tangibles, we augmented paper Pokémon cards with foil 36h11 AprilTags, $64 \times 64mm$ in size (Figure 3A and Figure 12 bottom row). The app recognizes the Pokémon once the card is placed onto the screen and slid into a designated spot (Figure 3B). In both applications, once an ID is established, it stays as metadata associated with the touch blob until the tangible is lifted from the screen.

Figure 3: AprilTag (36h11, feature size of 8mm) made of metal foil attached to a Pokémon card (A), which can be recognized by a companion app (B).

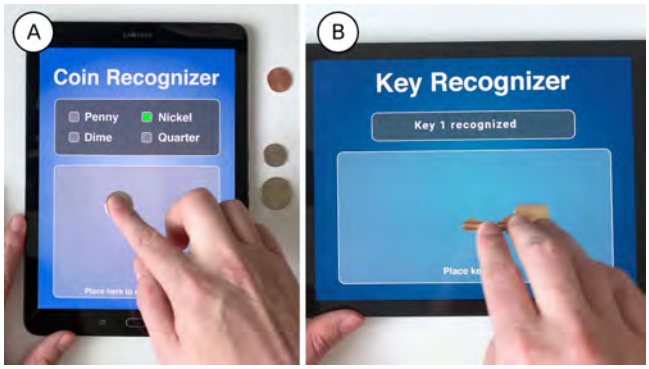

Figure 4: Super resolution can also be used to recognize everyday metal objects, such as coins (A) and keys (B).

4.2 Everyday Objects

In addition to special-purpose tangibles, described in the previous section, it could also be valuable for touchscreen devices to recognize and interact with unmodified everyday objects. In order to be sensed by the capacitive touchscreen, these objects must be made of a conductive material, such as metal or a carbon-loaded plastic.

To illustrate this possibility, we created a pair of similar applications, one that recognizes coins (Figure 4A) and another for keys (Figure 4B). The former could be part of educational software, teaching children to recognize different monetary values, while the latter could be a useful assistive feature.

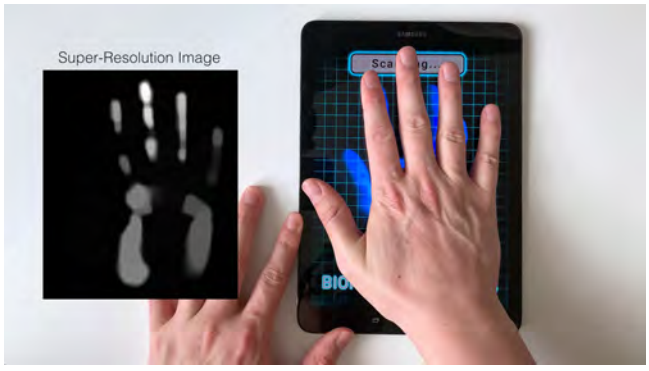

Figure 5: A super-resolution handprint scanner that could be used for user identification or authentication.

4.3 Users

Finally, touchscreens with super resolution could also be used to capture new dimensions of user input. For instance, using the shape and size of finger blobs in the capacitive touch image, Le et al. [33] demonstrated finger identification, while Mayer et al. [36] and Xiao et al. [60] estimated finger pitch and yaw. Without doubt, the accuracy of these prior systems could be boosted with the enhanced resolution of fingertips afforded by super resolution (see example super-resolution fingertips in Figure 15). Other prior work has used capacitive images to capture the shape of the ear [22] and palm [18] for identification and authentication, both of which could benefit from super resolution; see palm print example in Figures 2 and 5.

5 DATA COLLECTION

The major challenge in evaluating the quality of capacitive super-resolution output (in contrast to photography) is that high-resolution ground truth images do not exist. Instead, we use the real-world geometry of objects (measured with calipers) for evaluation.

As a very coarse grid of pixels, the impact of touchscreen quantization can vary dramatically depending on an object's position and orientation. Figure 6 illustrates the effect of quantization on a fine-grained feature at the limit of touchscreen sensor resolution ($4mm$). Figure 7 offers an example of how this effect manifests at object-scale. To investigate how the position and orientation of objects impacts the output quality, we captured data using a series of Movements with and without Rotation. More specifically, our Movements included left-right, up-down, circle, figure-eight, square, and random paths. We captured Rotation conditions for all non-round objects. For handprints, we rotated only as far as the touchscreen size would permit without any part of the hand leaving the screen.

5.1 Apparatus

For data collection, we use our aforementioned Samsung Galaxy Tab S2 tablet lying flat on a table. We prepared test sets for eight object Types, spanning different uses. To measure geometric accuracy,

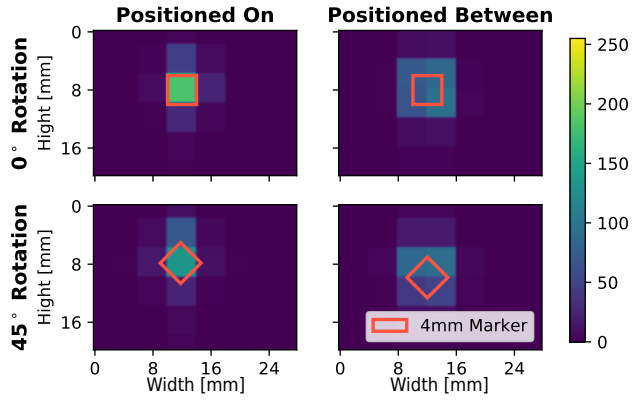

Figure 6: Signal of a $4 \times 4mm$ square in the capacitive image when positioned on or between capacitive pixels, as well as aligned vs. rotated $45°$.

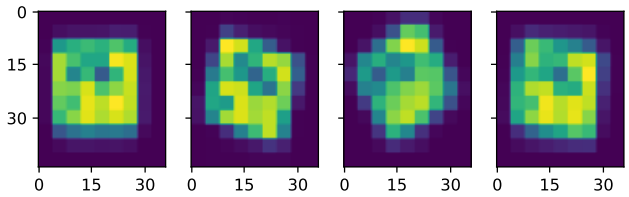

Figure 7: The same fiducial tag quantized at different translations and rotations on the screen. These are from the same frame sequence as the tag in Figure 12 (top row), and illustrates quality variance and the important of selecting a good seed frame.

we use: metal circles (diameter = 2, 4, 6, and 8mm), metal squares (side length = 2, 4, 6, and 8mm), and US coins (penny 19.05mm, nickel 21.21mm, dime 17.91mm, quarter 24.3mm). To investigate complex structured shapes, we made four low-density fiducials (feature diameter = 3, 4, 6, and 8mm) inspired by Yu et al. [67], and AprilTags (16h5 and 36h11 standards). Finally, to subjectively assess quality in resolving a complex object's shape, we used four keys, ten fingers, and two hands.

5.2 Procedure

For each object set Type, we recorded capacitive image data while systematically varying our independent variables (Movements and Rotation) in a nested design as rotation was not possible or needed in all combinations. For each of the 372 conditions, we recorded 30 seconds of data (at 16.2 FPS = ~500 samples). Additionally, we recorded 2 minutes of 9 different 36h11 AprilTags for the recognition evaluation. Thus, in total, we captured around 194K image frames. All parameters, including thresholds for sizing, were fixed before the study and applied uniformly to all test inputs.

5.3 Open Dataset

To bootstrap research in this area, and offer a benchmark for future evaluations, we have released all of our collected capacitive image data, which can be downloaded at https://github.com/FIGLAB/Super-Resolution-Dataset.git.

6 EVALUATION

To assess the geometric accuracy of our super-resolution pipeline, we used the data we captured for metal circles, metal squares, and US coins (4 Sizes each). These all have known, real-world sizes, from which we can compute the error. For fiducials, we test how well an off-the-shelf AprilTag recognizer [54] correct detects our super-resolved output.

6.1 Seed Selection

Our first analysis step was to determine how much variance in estimated size was generated by different seed images of an object, to which all other frames are aligned. For this analysis, we simply try all ~500 frames as seeds and look at the variance in output size. We found an average size variation of 0.41mm (SD = 0.45) across all tested objects, see Figure 8. The only unusual result was the US dime (17.91mm diameter), with an average size variation of

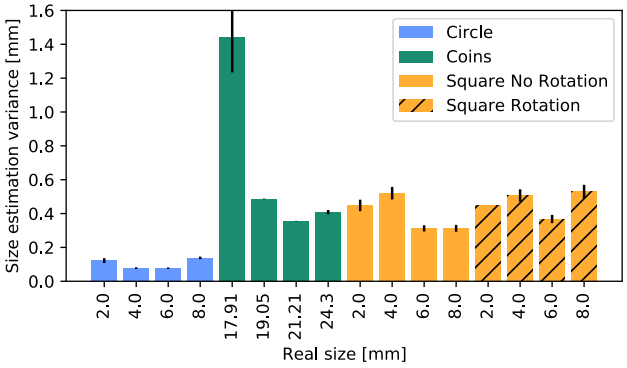

Figure 8: Average variance in estimated output size when trying all frames as seeds.

1.44mm across seed images; see Figure 8 left-most green bar. Thus, to achieve the very best results, we suggest adding a seed image selection stage before attempting super resolution. As noted in our Implementation section, we fit a bounding box to all available frames in a blob's time series and use the frame with the smallest bounding box as the seed image. We found this to be a simple, but reliable metric for good capacitive image pixel alignment (i.e., not unnecessarily crossing pixel boundaries).

6.2 Number of Frames vs. Quality

As our algorithm works with any sequence length of low-resolution input frames, we also wished to investigate how many frames are sufficient to generate a high-quality super-resolved output. For this, we ran a post hoc study, giving our pipeline different input sequence lengths, simulating different lengths of time, and using object size as our evaluation metric.

Figure 9 shows the accuracy over time for our four sizes of circles and squares. We found that quality converges quickly and that best quality is achieved within 35 to 45 images. This equates to ~0.6 seconds of data at the touch controller's native 60 FPS (or ~2.5 seconds of data with our prototype's 16.2 FPS capture rate). This latency may be acceptable for recognition of e.g., a tangible placed onto the screen, but would be unacceptable for conventional touch input. Thus, as noted previously, we envision a single-frame

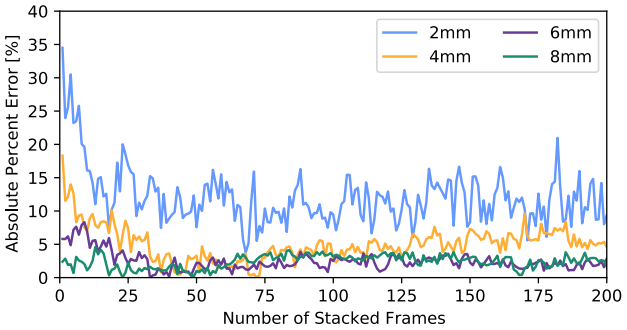

Figure 9: Geometric accuracy (normalized by object size) when running super-resolution on different input frame sequence lengths (i.e., simulating time).

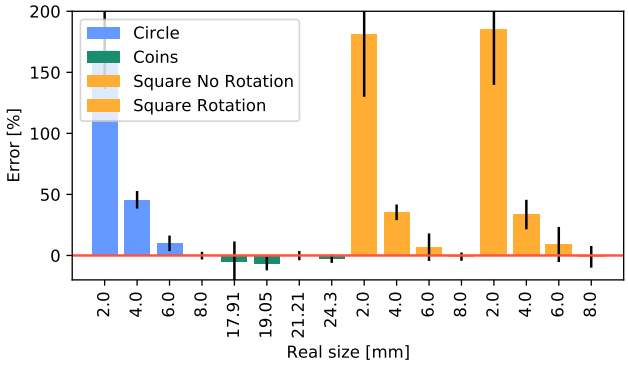

Figure 10: Geometric accuracy (normalized by object size) for all of our test circles, squares and coins (four sizes each). Square is additionally broken out by data with and without rotation (not applicable for our round objects).

touch tracking pipeline and multi-frame super resolution pipeline running in parallel, supporting different interactive needs. For all subsequent analyses, we only use frames 0 – 36 for super-resolution.

6.3 Accuracy Across Size and Rotation

Overall, we found that our super-resolution pipeline was able to resolve the size of objects to within 1.3mm (SD = 1.2). Figure 10 provides a full breakdown of these results. We note that all of our 2mm test objects were wildly inaccurate, showing the limits of what super-resolution can recover. If we exclude these three objects, mean size error is 0.8mm (SD = 0.6), or 12.5% (SD = 15.1) when normalized by object size.

For the squares, we performed a two-way Analysis of Variance (ANOVA), which revealed that there was a significant effect of Size on Accuracy ($F(3, 37) = 543.011$, $p < .001$, $\eta^2 = .978$). However, there was no significant effect of Rotation nor an interaction effect ($F(1, 37) = 0.085$, $p > .772$, $\eta^2 < .003$; ($F(3, 37) = 0.130$, p > .941, $\eta^2 < .011$, respectively). As Rotation has no effect on accuracy (see also square results in Figure 10), it suggests that our super-resolution implementation can properly handle object rotation and overcome the problems illustrated in Figures 6 and 7.

As we could not reveal an effect of Rotation for the squares, we combined the squares and circles and ran a one-way ANCOVA to see if there is a general effect of Size on accuracy with Type as a covariate. The ANCOVA revealed a statistically significant effect of Size on accuracy ($F(3, 4) = 309.311$, $p < .001$, $\eta^2 = .995$). We performed pairwise post hoc tests and found all $p < .001$ besides 6mm vs 8mm ($p = .200$). Further, we performed a one-way ANOVA with the four coins, confirming that larger elements get estimated more precisely as there was no statistically significant difference ($F(3, 20) = 2.591$, $p > .081$, $\eta^2 < .280$). This is not surprising as larger objects do cover more pixels; thus, more information is available to perform super-resolution.

6.4 Fiducials

To see how capacitive touchscreen super resolution works with complex structured patterns, we chose to use fiducial markers,

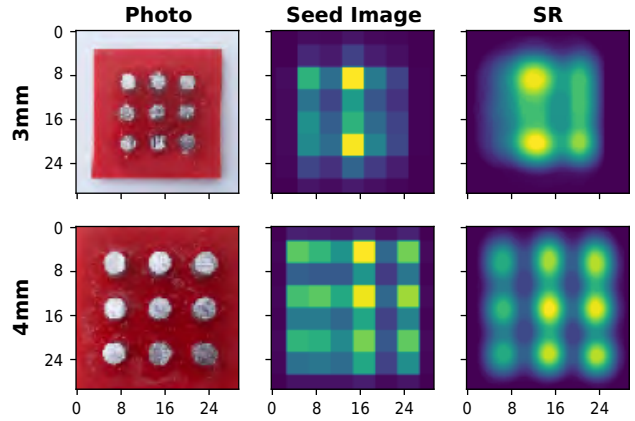

Figure 11: Examples of our 3×3 TUIC-inspired fiducials; top row uses pads 3mm in diameter, while the bottom row uses 4mm pads. Left: photo reference. Center: best seed image. Right: super-resolved fiducial.

which could be used to support libraries of tangibles. We tested two designs: low-density fiducial markers (similar to TUIC [67]) and standard AprilTags [54].

Our first fiducial set was inspired by those used in TUIC [67] – a 3×3 grid of circular metal pads (one for each bit). We created four different tags – 3, 4, 6, and 8mm – which denote both the size of pads and space between the pads; see Figure 11 for an example. As we found in our evaluation, pad size/spacing of 6 and 8mm were easily resolved. Tags of 4mm size/spacing were generally able to be recognized (Figure 11, bottom row), but 3mm tags almost always failed (Figure 11, top row). As a point of reference, our 3×3 tag with 4mm pads was $20 \times 20mm$ ($44mm^2$ per bit), while TUIC's [67] 5×5-pad tag was $50 \times 50mm$ ($100mm^2$ per bit).

Our second fiducial set consists of standard AprilTags [54] made of aluminum foil. These can be recognized by an unmodified AprilTag reader. We created $64 \times 64mm$ tags using the 36h11 AprilTags format (36-bit raw payload; 587 unique IDs with error correction). The bottom row of Figure 12 shows an example tag at its various processing stages. We also created several smaller 16h5 AprilTags

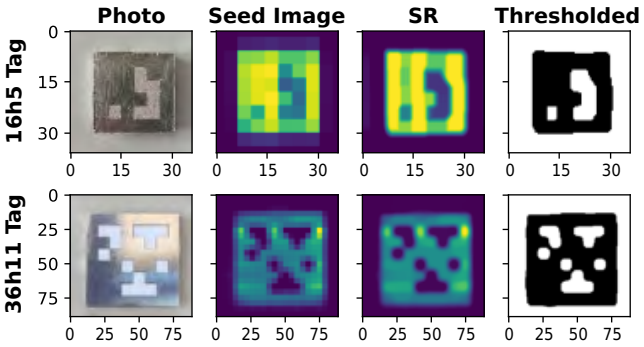

Figure 12: Examples of a $24 \times 24mm$ 16h5 AprilTag (top row) and $64 \times 64mm$ 36h11 AprilTag (bottom row) at different processing stages.

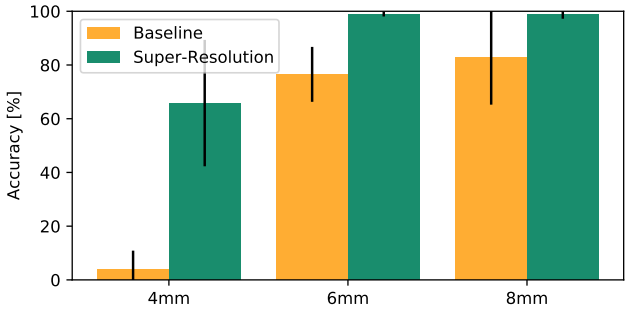

Figure 13: Recognition accuracy of capacitive AprilTags (36h11) with features sizes of 4, 6, and 8mm, across 3137 instances using capacitive images (baseline) and super-resolved output.

for testing and found that $24 \times 24mm$ tags worked well (Figure 12, top row; $36mm^2$ per bit). These are small enough to be placed under, for example, chess pieces and other small tangibles, see Figure 1.

To evaluate recognition accuracy, we tested three 36h11 AprilTags (ID 2, 3, and 4) in three different sizes (feature sizes of 4, 6, and 8mm). We recorded capacitive data from all nine fiducial markers for two minutes each, totaling around 16k samples. For the baseline, we normalized and thresholded the image to help the recognition. When using the raw capacitive image, fiducial recognition was 4.2%, 76.5% and 82.5% accurate for the AprilTags with 4, 6, and 8mm features, respectively. With super-resolution, recognition accuracy jumps to 66.2%, 99.0% and 98.8%, respectively. A related t-test ($t(18) = -3.973$, $p < .005$) showed that our super-resolution method ($M = 87.9\%$ $SD = 32.6\%$) significantly outperforms the baseline ($M = 54.2\%$ $SD = 49.8$).

7 OTHER POTENTIAL USES

We also captured data for four door keys as another example class of complex objects (Figure 14). We did not perform any quantitative analysis; however, we prototyped a demo use-case app, see Figure 14. Note this would not work for all key types, such as dimple keys.

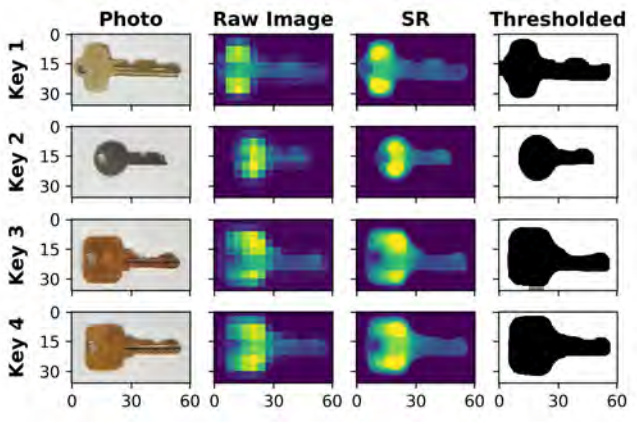

Figure 14: Four keys at different stages of processing.

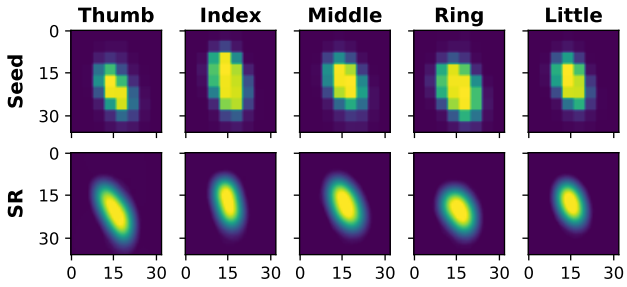

Figure 15: Low resolution input (seed image) and super-resolution output of five example fingertips.

Fiducials are a good exemplar of a man-made, geometrically-complex input. As a compliment to this data, we also captured human input examples, in the form of fingers and hands, to study natural shapes. We did not perform any formal experimentation on this data, as biometric analyses are a specialized domain beyond the scope of our investigations. However, we did run all data through our pipeline, with example output shown in Figures 2 and 15.

8 LIMITATIONS

It is important to note that our method would not impact the latency of conventional touch events (e.g., pressing a close button, typing on a soft keyboard), as we envision our super-resolution pipeline running in parallel with standard single-frame touch pipelines. Our method's inherent need for many capacitive image frames does mean it operates with higher latency. Fortunately, input events such as placing a tangible down onto a screen or scanning one's palm for identification are much less latency-sensitive than traditional finger touch events. Nonetheless, thoughtful application of when to leverage super-resolved touchscreen input will be required to minimize this limitation.

The low resolution of capacitive touchscreens is a second challenge. These commodity components are value-engineered to be good at capturing fingers, and little else, setting a lower bound capacitive pixel size of around 4mm. This pitch ensures fingers are seen by at least a few pixels, permitting good interpolation of the touch centroid (e.g., [64]) but is very poor at capturing any geometric detail. A severe consequence of this low sensing resolution is very coarse alignment and stacking of object frames, which affects all downstream processing, reducing the quality of super-resolved output. This limitation is most severe for smaller objects, where there are fewer pixels and few unique geometric details on which to align frames (e.g., coins, finger-tips). For larger and more complex items (e.g., fiducials, handprints), the effect is lessened.

A third limitation of our approach is the need for objects (e.g., fingers, tangibles) to be translated across the surface of the screen. This is the only way to quantize the object at a variety of alignments, which allows super resolution to resolve sub-pixel details. However, as we noted above, the quality follows a logarithmic trend with respect to the number of frames captured. In our experience, we found that acceptable quality (e.g., correct AprilTag recognition) was typically achieved within 10 frames, or roughly half a second (see also Video Figure for real-time examples of this effect). This means, in practice, objects have to be translated for at least a sub-pixel distance over the screen (i.e., ~4mm). In fact, movements smaller than a capacitive pixel are better for the super-resolution process as the quantization pattern repeats every pixel; and thus, adds very little to no information. Nonetheless, some movement is necessary, and it means we cannot apply super-resolution to e.g., finger taps or tangibles placed "perfectly" down onto the screen (without any movement or jitter).

Finally, our proof-of-concept implementation runs on a 2017 MacBook Pro and is computationally taxing even on this hardware. Thus, future work remains to build a self-contained smartphone version. That said, we implemented most of our software stack in python (chiefly to facilitate rapid prototyping), which is not known for its high performance. Moving to a compiled and hardware-accelerated implementation should yield significant performance gains, able to run on smartphone hardware.

9 CONCLUSION

We have presented our work on bringing super resolution techniques, long used in other domains, to capacitive touchscreens. We show that through software alone, we can increase the resolution of touchscreen sensors. This extra fidelity allows us to e.g., support large libraries of tangibles though higher-information-density capacitive fiducials than previously demonstrated. Our evaluations and example apps show how everyday objects, such as coins and keys can be recognized for assistive and educational applications. While we view our work as a useful first step, demonstrating proof of concept, we believe more gains are likely in the future with more advanced techniques.

REFERENCES

[1] Simon Baker and Takeo Kanade. 2002. Limits on super-resolution and how to break them. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 9 (2002), 1167–1183. https://doi.org/10.1109/TPAMI.2002.1033210

[2] Eric Betzig, George H. Patterson, Rachid Sougrat, O. Wolf Lindwasser, Scott Olenych, Juan S. Bonifacino, Michael W. Davidson, Jennifer Lippincott-Schwartz, and Harald F. Hess. 2006. Imaging Intracellular Fluorescent Proteins at Nanometer Resolution. *Science* 313, 5793 (2006), 1642–1645. https://doi.org/10.1126/science.1127344

[3] David S Bolme, J Ross Beveridge, Bruce A Draper, and Yui Man Lui. 2010. Visual object tracking using adaptive correlation filters. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*. IEEE, 2544–2550. https://doi.org/10.1109/CVPR.2010.5539960

[4] Roberto Brunelli. 2009. *Template Matching Techniques in Computer Vision: Theory and Practice*. John Wiley & Sons.

[5] Liwei Chan, Stefanie Müller, Anne Roudaut, and Patrick Baudisch. 2012. CapStones and ZebraWidgets: Sensing Stacks of Building Blocks, Dials and Sliders on Capacitive Touch Screens. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 2189–2192. https://doi.org/10.1145/2207676.2208371

[6] Peter Dalsgaard and Kim Halskov. 2012. Tangible 3D Tabletops: Combining Tangible Tabletop Interaction and 3D Projection. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design* (Copenhagen, Denmark) *(NordiCHI '12)*. Association for Computing Machinery, New York, NY, USA, 109–118. https://doi.org/10.1145/2399016.2399033

[7] Paul Dietz and Darren Leigh. 2001. DiamondTouch: A Multi-User Touch Technology. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology* (Orlando, Florida) *(UIST '01)*. Association for Computing Machinery, New York, NY, USA, 219–226. https://doi.org/10.1145/502348.502389

[8] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. 2016. Image Super-Resolution Using Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38, 2 (2016), 295–307. https://doi.org/10.1109/TPAMI.2015.2439281

[9] Thibaud Ehret, Axel Davy, Pablo Arias, and Gabriele Facciolo. 2019. Joint Demosaicking and Denoising by Fine-Tuning of Bursts of Raw Images. In

IEEE/CVF International Conference on Computer Vision (ICCV '19). IEEE, 8867–8876. https://doi.org/10.1109/ICCV.2019.00896

[10] Georgios D. Evangelidis and Emmanouil Z. Psarakis. 2008. Parametric Image Alignment Using Enhanced Correlation Coefficient Maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30, 10 (2008), 1858–1865. https://doi.org/10.1109/TPAMI.2008.113

[11] James R. Fienup. Fienup. 1997. Invariant error metrics for image reconstruction. *Applied optics* 36, 32 (Nov 1997), 8352–8357. https://doi.org/10.1364/AO.36.008352

[12] Andrew W. Fitzgibbon and Robert B. Fisher. 1995. A Buyer's Guide to Conic Fitting. In *Proceedings of the 6th British conference on Machine vision (BMVC '95, Vol. 2)*. BMVA Press, 513–522. https://doi.org/10.5244/C.9.51

[13] George W. Fitzmaurice, Hiroshi Ishii, and William A. S. Buxton. 1995. Bricks: Laying the Foundations for Graspable User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '95)*. ACM Press/Addison-Wesley Publishing Co., USA, 442–449. https://doi.org/10.1145/223904.223964

[14] Tom Goldstein and Stanley Osher. 2009. The split Bregman method for L1-regularized problems. *SIAM Journal on Imaging Sciences 2* 2 (2009), 323–343. https://doi.org/10.1137/080725891

[15] Hayit Greenspan. 2008. Super-Resolution in Medical Imaging. *Comput. J.* 52, 1 (02 2008), 43–63. https://doi.org/10.1093/comjnl/bxm075

[16] Tobias Grosse-Puppendahl, Christian Holz, Gabe Cohn, Raphael Wimmer, Oskar Bechtold, Steve Hodges, Matthew S. Reynolds, and Joshua R. Smith. 2017. Finding Common Ground: A Survey of Capacitive Sensing in Human-Computer Interaction. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3293–3315. https://doi.org/10.1145/3025453.3025808

[17] Manuel Guizar-Sicairos, Samuel T. Thurman, and James R. Fienup. 2008. Efficient subpixel image registration algorithms. *Opt. Lett.* 33, 2 (Jan 2008), 156–158. https://doi.org/10.1364/OL.33.000156

[18] Anhong Guo, Robert Xiao, and Chris Harrison. 2015. CapAuth: Identifying and Differentiating User Handprints on Commodity Capacitive Touchscreens. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces* (Madeira, Portugal) *(ITS '15)*. Association for Computing Machinery, New York, NY, USA, 59–62. https://doi.org/10.1145/2817721.2817722

[19] Nils Gustafsson, Siân Culley, George Ashdown, Dylan M. Owen, Pedro Matos Pereira, and Ricardo Henriques. 2016. Fast live-cell conventional fluorophore nanoscopy with ImageJ through super-resolution radial fluctuations. *Nature Communications* 7, 1 (12 Aug 2016), 12471. https://doi.org/10.1038/ncomms12471

[20] Changyo Han, Katsufumi Matsui, and Takeshi Naemura. 2020. ForceStamps: Fiducial Markers for Pressure-Sensitive Touch Surfaces to Support Rapid Prototyping of Physical Control Interfaces. In *Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction* (Sydney NSW, Australia) *(TEI '20)*. Association for Computing Machinery, New York, NY, USA, 273–285. https://doi.org/10.1145/3374920.3374924

[21] David Held, Sebastian Thrun, and Silvio Savarese. 2016. Learning to Track at 100 FPS with Deep Regression Networks. In *Computer Vision – ECCV 2016 (ECCV '16)*, Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling (Eds.). Springer International Publishing, Cham, 749–765. https://doi.org/10.1007/978-3-319-46448-0_45 arXiv:1604.01802

[22] Christian Holz, Senaka Buthpitiya, and Marius Knaust. 2015. Bodyprint: Biometric User Identification on Mobile Devices Using the Capacitive Touchscreen to Scan Body Parts. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3011–3014. https://doi.org/10.1145/2702123.2702518

[23] Mareki Honma, Kazunori Akiyama, Makoto Uemura, and Shiro Ikeda. 2014. Super-resolution imaging with radio interferometry using sparse modeling. *Publications of the Astronomical Society of Japan* 66, 5 (09 2014), 14. https://doi.org/10.1093/pasj/psu070 95.

[24] Hsieh Hou and H. Andrews. 1978. Cubic splines for image interpolation and digital filtering. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 26, 6 (1978), 508–517. https://doi.org/10.1109/TASSP.1978.1163154

[25] Hiroshi Ishii and Brygg Ullmer. 1997. Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '97)*. Association for Computing Machinery, New York, NY, USA, 234–241. https://doi.org/10.1145/258549.258715

[26] Zdenek Kalal, Krystian Mikolajczyk, and Jiri Matas. 2010. Forward-Backward Error: Automatic Detection of Tracking Failures. In *Proceedings of the 20th International Conference on Pattern Recognition (ICPR '10)*. IEEE, 2756–2759. https://doi.org/10.1109/ICPR.2010.675

[27] Martin Kaltenbrunner and Ross Bencina. 2007. ReacTIVision: A Computer-Vision Framework for Table-Based Tangible Interaction. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction* (Baton Rouge, Louisiana) *(TEI '07)*. Association for Computing Machinery, New York, NY, USA, 69–74. https://doi.org/10.1145/1226969.1226983

[28] Levon Khachatryan. 2019. Centroid Based Object Tracking. https://github.com/lev1khachatryan/Centroid-Based_Object_Tracking.

[29] Sven Kratz, Tilo Westermann, Michael Rohs, and Georg Essl. 2011. CapWidgets: Tangile Widgets versus Multi-Touch Controls on Mobile Devices. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI EA '11)*. Association for Computing Machinery, New York, NY, USA, 1351–1356. https://doi.org/10.1145/1979742.1979773

[30] Abinaya Kumar, Aishwarya Radjesh, Sven Mayer, and Huy Viet Le. 2019. Improving the Input Accuracy of Touchscreens Using Deep Learning. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3290607.3312928

[31] Huy Viet Le, Thomas Kosch, Patrick Bader, Sven Mayer, and Niels Henze. 2018. PalmTouch: Using the Palm as an Additional Input Modality on Commodity Smartphones. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173934

[32] Huy Viet Le, Sven Mayer, and Niels Henze. 2018. InfiniTouch: Finger-Aware Interaction on Fully Touch Sensitive Smartphones. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (Berlin, Germany) *(UIST '18)*. Association for Computing Machinery, New York, NY, USA, 779–792. https://doi.org/10.1145/3242587.3242605

[33] Huy Viet Le, Sven Mayer, and Niels Henze. 2019. Investigating the Feasibility of Finger Identification on Capacitive Touchscreens Using Deep Learning. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (Marina del Ray, California) *(IUI '19)*. Association for Computing Machinery, New York, NY, USA, 637–649. https://doi.org/10.1145/3301275.3302295

[34] Rong-Hao Liang, Han-Chih Kuo, Liwei Chan, De-Nian Yang, and Bing-Yu Chen. 2014. GaussStones: Shielded Magnetic Tangibles for Multi-Token Interactions on Portable Displays. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) *(UIST '14)*. Association for Computing Machinery, New York, NY, USA, 365–372. https://doi.org/10.1145/2642918.2647384

[35] Zhouchen Lin and Heung-Yeung Shum. 2004. Fundamental limits of reconstruction-based superresolution algorithms under local translation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26, 1 (2004), 83–97. https://doi.org/10.1109/TPAMI.2004.1261081

[36] Sven Mayer, Huy Viet Le, and Niels Henze. 2017. Estimating the Finger Orientation on Capacitive Touchscreens Using Convolutional Neural Networks. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces* (Brighton, United Kingdom) *(ISS '17)*. Association for Computing Machinery, New York, NY, USA, 220–229. https://doi.org/10.1145/3132272.3134130

[37] Microsoft Corporation. 2008 (accessed December 13, 2020). *PixelSense*. Microsoft Corporation. https://en.wikipedia.org/wiki/Microsoft_PixelSense

[38] Viet Nguyen, Siddharth Rupavatharam, Luyang Liu, Richard Howard, and Marco Gruteser. 2019. HandSense: Capacitive Coupling-Based Dynamic, Micro Finger Gesture Recognition. In *Proceedings of the 17th Conference on Embedded Networked Sensor Systems* (New York, New York) *(SenSys '19)*. Association for Computing Machinery, New York, NY, USA, 285–297. https://doi.org/10.1145/3356250.3360040

[39] Alex Olwal and Andrew D. Wilson. 2008. SurfaceFusion: Unobtrusive Tracking of Everyday Objects in Tangible User Interfaces. In *Proceedings of Graphics Interface 2008* (Windsor, Ontario, Canada) *(GI '08)*. Canadian Information Processing Society, CAN, 235–242.

[40] Esben Warming Pedersen and Kasper Hornbæk. 2011. Tangible Bots: Interaction with Active Tangibles in Tabletop Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 2975–2984. https://doi.org/10.1145/1978942.1979384

[41] Klaus G. Puschmann and Franz Kneer. 2005. On super-resolution in astronomical imaging. *Astronomy & Astrophysics* 436, 1 (2005), 373–378. https://doi.org/10.1051/0004-6361:20042320

[42] Jun Rekimoto. 2002. SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Minneapolis, Minnesota, USA) *(CHI '02)*. Association for Computing Machinery, New York, NY, USA, 113–120. https://doi.org/10.1145/503376.503397

[43] Adrian Rosebrock. 2018. Simple object tracking with OpenCV. https://www.pyimagesearch.com/2018/07/23/simple-object-tracking-with-opencv/.

[44] Martin Schmitz, Jürgen Steimle, Jochen Huber, Niloofar Dezfuli, and Max Mühlhäuser. 2017. Flexibles: Deformation-Aware 3D-Printed Tangibles for Capacitive Touchscreens. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 1001–1014. https://doi.org/10.1145/3025453.3025663

[45] Jack Sklansky. 1982. Finding the Convex Hull of a Simple Polygon. *Pattern Recogn. Lett.* 1, 2 (Dec. 1982), 79–83. https://doi.org/10.1016/0167-8655(82)90016-2

[46] Satoshi Suzuki and Keiicji Abe. 1985. Topological structural analysis of digitized binary images by border following. *Computer Vision, Graphics, and Image Processing* 30, 1 (1985), 32 – 46. https://doi.org/10.1016/0734-189X(85)90016-7

[47] Godfried T. Toussaint. 1983. Solving geometric problems with the rotating calipers. In *Proc. IEEE Melecon*, Vol. 83. IEEE, A10.

[48] Brygg Ullmer and Hiroshi Ishii. 1997. The MetaDESK: Models and Prototypes for Tangible User Interfaces. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology* (Banff, Alberta, Canada) *(UIST '97)*. Association for Computing Machinery, New York, NY, USA, 223–232. https://doi.org/10.1145/263407.263551

[49] John Underkoffler and Hiroshi Ishii. 1999. Urp: A Luminous-Tangible Workbench for Urban Planning and Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Pittsburgh, Pennsylvania, USA) *(CHI '99)*. Association for Computing Machinery, New York, NY, USA, 386–393. https://doi.org/10.1145/302979.303114

[50] Nicolas Villar, Daniel Cletheroe, Greg Saul, Christian Holz, Tim Regan, Oscar Salandin, Misha Sra, Hui-Shyong Yeo, William Field, and Haiyan Zhang. 2018. Project Zanzibar: A Portable and Flexible Tangible Interaction Platform. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3174089

[51] Simon Voelker, Christian Cherek, Jan Thar, Thorsten Karrer, Christian Thoresen, Kjell Ivar Øvergård, and Jan Borchers. 2015. PERCs: Persistently Trackable Tangibles on Capacitive Multi-Touch Displays. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (Charlotte, NC, USA) *(UIST '15)*. Association for Computing Machinery, New York, NY, USA, 351–356. https://doi.org/10.1145/2807442.2807466

[52] Simon Voelker, Kosuke Nakajima, Christian Thoresen, Yuichi Itoh, Kjell Ivar Øvergård, and Jan Borchers. 2013. PUCs: Detecting Transparent, Passive Untouched Capacitive Widgets on Unmodified Multi-Touch Displays. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces* (St. Andrews, Scotland, United Kingdom) *(ITS '13)*. Association for Computing Machinery, New York, NY, USA, 101–104. https://doi.org/10.1145/2512349.2512791

[53] Geoff Walker. 2012. A review of technologies for sensing contact location on the surface of a display. *Journal of the Society for Information Display* 20, 8 (2012), 413–440. https://doi.org/10.1002/jsid.100

[54] John Wang and Edwin Olson. 2016. AprilTag 2: Efficient and robust fiducial detection. In *RSJ International Conference on Intelligent Robots and Systems (IROS) (IROS '16)*. IEEE, 4193–4198. https://doi.org/10.1109/IROS.2016.7759617

[55] Yilun Wang, Junfeng Yang, Wotao Yin, and Yin Zhang. 2008. A new alternating minimization algorithm for total variation image reconstruction. *SIAM Journal on Imaging Sciences* 1, 3 (2008), 248–272. https://doi.org/10.1137/080724265

[56] Frederick W. Wheeler, Xiaoming Liu, and Peter H. Tu. 2007. Multi-Frame Super-Resolution for Face Recognition. In *First IEEE International Conference on Biometrics: Theory, Applications, and Systems*. IEEE, 1–6. https://doi.org/10.1109/BTAS.2007.4401949

[57] Andrew D. Wilson and Raman Sarin. 2007. BlueTable: Connecting Wireless Mobile Devices on Interactive Surfaces Using Vision-Based Handshaking. In *Proceedings of Graphics Interface 2007* (Montreal, Canada) *(GI '07)*. Association for Computing Machinery, New York, NY, USA, 119–125. https://doi.org/10.1145/1268517.1268539

[58] Bartlomiej Wronski, Ignacio Garcia-Dorado, Manfred Ernst, Damien Kelly, Michael Krainin, Chia-Kai Liang, Marc Levoy, and Peyman Milanfar. 2019. Handheld Multi-Frame Super-Resolution. *ACM Trans. Graph.* 38, 4, Article 28 (July 2019), 18 pages. https://doi.org/10.1145/3306346.3323024

[59] Robert Xiao, Scott Hudson, and Chris Harrison. 2016. CapCam: Enabling Rapid, Ad-Hoc, Position-Tracked Interactions Between Devices. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces* (Niagara Falls, Ontario, Canada) *(ISS '16)*. Association for Computing Machinery, New York, NY, USA, 169–178. https://doi.org/10.1145/2992154.2992182

[60] Robert Xiao, Julia Schwarz, and Chris Harrison. 2015. Estimating 3D Finger Angle on Commodity Touchscreens. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces* (Madeira, Portugal) *(ITS '15)*. Association for Computing Machinery, New York, NY, USA, 47–50. https://doi.org/10.1145/2817721.2817737

[61] Li Xu, Cewu Lu, Yi Xu, and Jiaya Jia. 2011. Image Smoothing via L0 Gradient Minimization. In *Proceedings of the 2011 SIGGRAPH Asia Conference* (Hong Kong, China) *(SA '11)*. Association for Computing Machinery, New York, NY, USA, Article 174, 12 pages. https://doi.org/10.1145/2024156.2024208

[62] Li Xu, Shicheng Zheng, and Jiaya Jia. 2013. Unnatural L0 Sparse Representation for Natural Image Deblurring. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '13)*. IEEE, 1107–1114. https://doi.org/10.1109/CVPR.2013.147

[63] Xiangyu Xu, Yongrui Ma, and Wenxiu Sun. 2019. Towards Real Scene Super-Resolution With Raw Images. In *Conference on Computer Vision and Pattern Recognition (CVPR '19)*. IEEE, 1723–1731. https://doi.org/10.1109/CVPR.2019.00182

[64] Mehrdad Yaghoobi, Stephen McLaughlin, and Mike E Davies. 2013. Super-resolution Sparse Projected Capacitive Multitouch Sensing. *IET Conference Proceedings*, 1.3–1.3(1). https://doi.org/10.1049/cp.2013.2041

[65] Jianchao Yang, John Wright, Thomas S. Huang, and Yi Ma. 2010. Image Super-Resolution Via Sparse Representation. *IEEE Transactions on Image Processing* 19, 11 (2010), 2861–2873. https://doi.org/10.1109/TIP.2010.2050625

[66] Alper Yilmaz, Omar Javed, and Mubarak Shah. 2006. Object Tracking: A Survey. *Comput. Surveys* 38, 4 (Dec. 2006), 13–es. https://doi.org/10.1145/1177352.1177355

[67] Neng-Hao Yu, Li-Wei Chan, Seng Yong Lau, Sung-Sheng Tsai, I-Chun Hsiao, Dian-Je Tsai, Fang-I Hsiao, Lung-Pan Cheng, Mike Chen, Polly Huang, and Yi-Ping Hung. 2011. TUIC: Enabling Tangible Interaction on Capacitive Multi-Touch Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 2995–3004. https://doi.org/10.1145/1978942.1979386

[68] Yang Zhang, Wolf Kienzle, Yanjun Ma, Shiu S. Ng, Hrvoje Benko, and Chris Harrison. 2019. ActiTouch: Robust Touch Detection for On-Skin AR/VR Interfaces. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 1151–1159. https://doi.org/10.1145/3332165.3347869