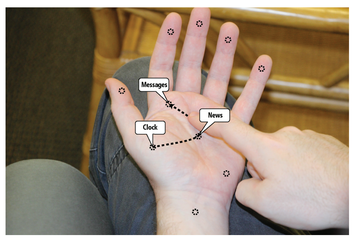
Localization of Skin Features on the Hand and Wrist from Small Image Patches

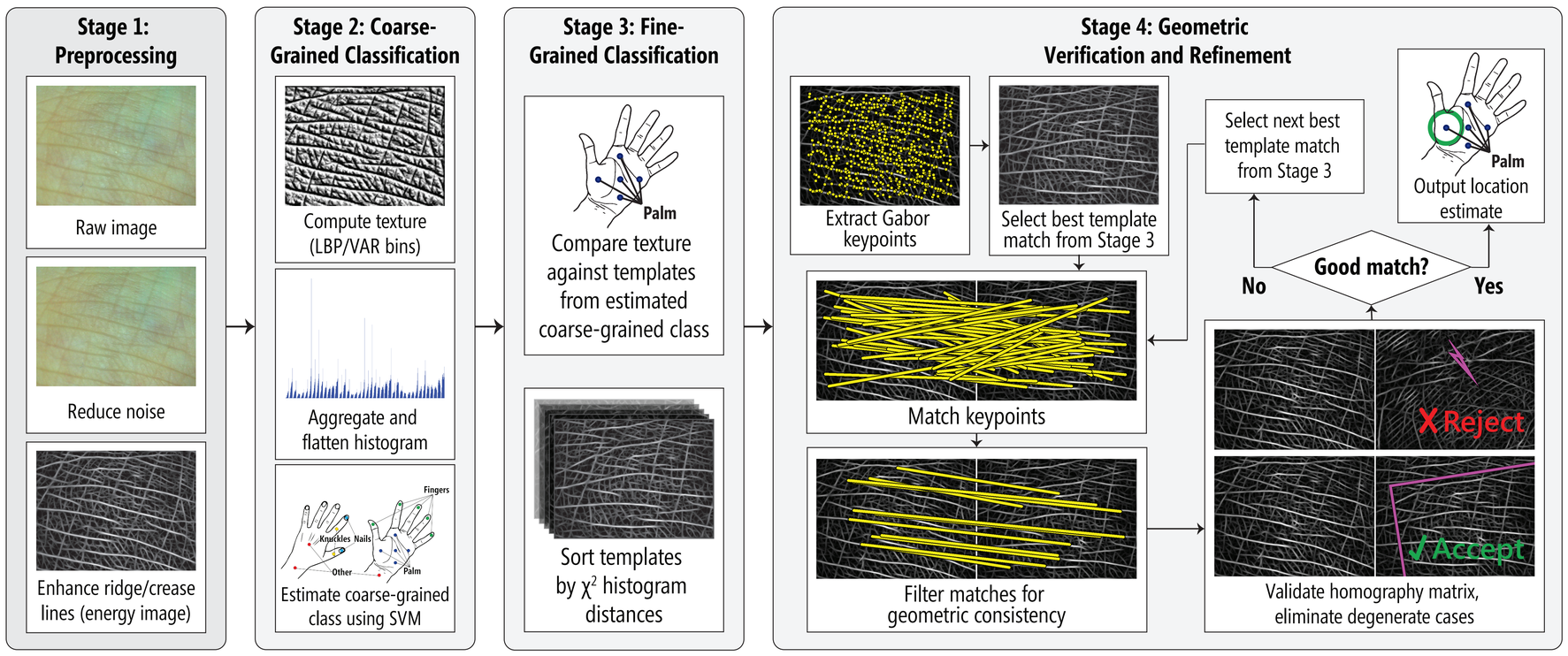
Skin-based biometrics rely on the distinctiveness of skin patterns across individuals for identification. In contrast, our research investigates whether small image patches of the skin can be localized on a user's body, determining not ***"who?"***, but instead ***"where?"***

This capability could enable a wearable camera to support a range of ***on-body interactions***, where users tap or gesture on their own body to control a computing device. One advantage of on-body input is that it is always available, allowing the user to, for example, quickly tap or swipe on their palm to answer a phone call or listen to new emails. It is also useful when visual attention is limited because the skin's tactile perception allows for more accurate input than is possible with a touchscreen.



Conceptual visualization of on-hand input to control a mobile device.

Touch Localization Pipeline



We estimate the user's touch location from close-up images using a four-stage hierarchical classifier. Shown above is an example from the left side of the palm as each of the four stages are applied.

Stage 1: Preprocessing. First, the image is preprocessed to remove surface artifacts and camera noise, and to emphasize ridge and crease lines using the energy image (defined at each pixel as the maximum Gabor response).

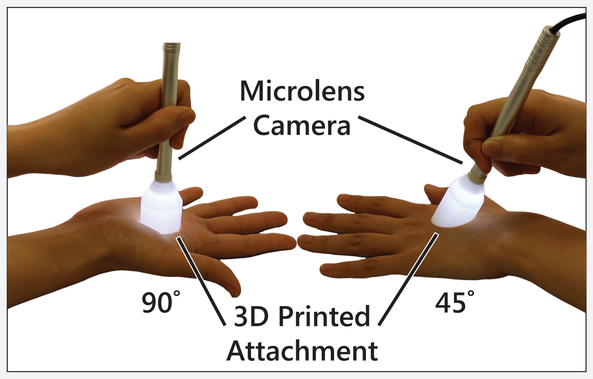
Stage 2: Coarse-Grained Classification. Second, the image's texture, defined as the 2D histogram of LBP and pixel variances, is classified into one of five coarse-grained locations (in this case, the palm).

Stage 3: Fine-Grained Classification. Third, the image's texture is compared against the templates from the predicted coarse-grained class, which are sorted by their χ^2 histogram distances to prioritize matching for the next stage.

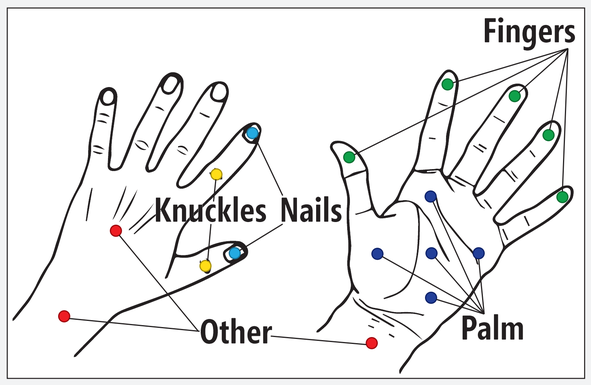
Stage 4: Geometric Verification and Refinement. Finally, the image is compared against the individual training templates, using a set of custom Gabor keypoints and descriptors. If a geometrically consistent match is found, then the fine-grained location can be estimated with a high degree of certainty (in this case, the left side of the palm); otherwise, the algorithm falls back upon the closest texture match from Stage 3.

Dataset

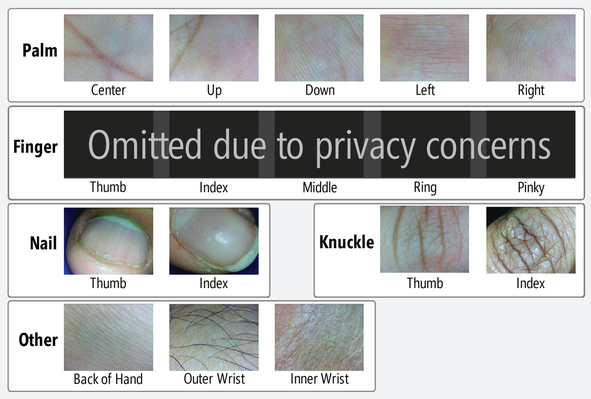
We created an image dataset from 30 volunteers across a variety of demographics. For each participant, we captured 20 images at each of 17 locations for a total of 10,198 images (one participant skipped two trials).



To capture close-up images of the skin, we used a small, self-illuminated 0.3MP camera with two custom 3D-printed attachments to control for distance and perspective. We expect future iterations to be much smaller.



Our dataset consists of 17 fine-grained locations on the left hand and wrist, grouped into 5 coarse-grained locations (color coded above).



Representative images from our dataset, selected across 12 participants. While we would like to release this dataset publicly, we cannot do so without risking the privacy of our participants.

Experiments and Results

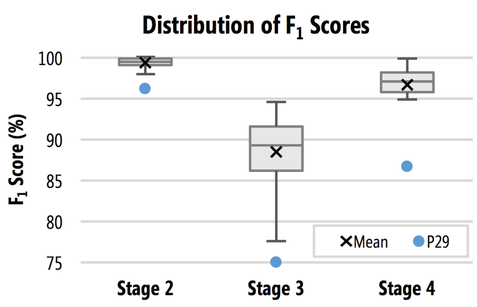
Within-Person Classification

To evaluate the overall location-level classifiability of the hand, we conducted a within-person classification experiment. Our strong performance suggests that skin-surface image patches can be classified and localized on the body with high levels of accuracy.

Stage 2: Coarse-grained texture (5 classes):
precision=99.1% (*SD*=0.9%), recall=99.2% (*SD*=0.8%)

Stage 3: Fine-grained texture (17 classes):
precision=88.2% (*SD*=4.4%), recall=88.0% (*SD*=4.5%)

Stage 4: Geometric feature verification (17 classes):
precision=96.6% (SD=2.2%), recall=96.4% (SD=2.3%)



Distribution of F_1 scores by participant, with outlier P29 (blue dot). We found that the top performing participants had high-contrast skin textures, more consistent pressure (resulting in fewer variations in lighting and focus), and greater consistency in returning to the same touch location.

Stage 2: Coarse-grained Classification Confusion Matrix

	Palm	*Finger*	*Nail*	*Knuckle*	*Other*
Palm	**99.0%**	0.5%			0.5%
Finger	0.6%	**99.3%**	0.1%		
Nail	0.2%	0.1%	**99.7%**	0.1%	
Knuckle			0.2%	**99.1%**	0.7%
Other	0.6%		0.1%	0.5%	**98.8%**

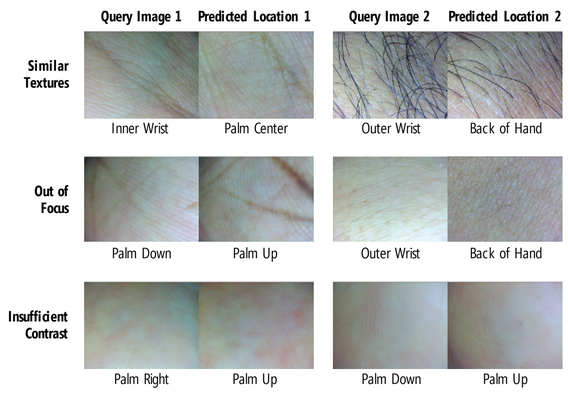
Classification percentages for classes at the coarse-grained level.

Stage 4: Fine-grained Classification Confusion Matrix

	Palm					*Fingers*					*Nails*		*Knuckles*		*Other*		
	C	*U*	*D*	*L*	*R*	*1st*	*2nd*	*3rd*	*4th*	*5th*	*1st*	*2nd*	*1st*	*2nd*	*BH*	*OW*	*IW*
Palm Center (C)	**98.3%**					0.2%		0.2%		0.2%					0.2%	0.5%	0.5%
Palm Up (U)	0.2%	**98.5%**		0.2%	0.2%		0.2%			0.3%						0.2%	0.3%
Palm Down (D)	0.3%	1.2%	**95.7%**	0.2%	1.7%	0.3%	0.2%						0.2%			0.2%	0.2%
Palm Left (L)	0.3%	0.3%	0.2%	**98.7%**	0.3%	0.3%											0.7%
Palm Right (R)	0.7%	0.5%	0.3%	0.5%	**97.5%**	0.2%	0.2%			0.2%							
1st Finger		0.5%	0.2%	0.2%	0.7%	**96.3%**	0.3%	0.5%	0.5%	0.7%					0.2%		
2nd Finger		0.3%			0.2%	0.3%	**95.8%**	1.7%	0.5%	1.2%							
3rd Finger			0.3%			0.2%	1.3%	**95.4%**	2.2%	0.7%							
4th Finger			0.2%				0.3%	1.8%	**95.3%**	2.3%							
5th Finger		0.2%		0.2%			0.3%	0.5%	1.5%	**97.0%**	0.3%						
1st Nail				0.2%							**98.2%**	1.7%					
2nd Nail			02%						0.2%		0.5%	**99.0%**		0.2%			
1st Knuckle											0.2%		**97.3%**	1.2%	0.2%	0.2%	1.0%
2nd Knuckle												0.2%	0.8%	**98.8%**		0.2%	
Back of Hand (BH)				0.2%									0.5%	0.2%	**92.2%**	4.7%	2.3%
Outer Wrist (OW)	0.2%												0.2%		6.0%	**90.2%**	3.5%
Inner Wrist (IW)	0.7%	0.7%	0.2%	0.2%	0.3%						0.2%		0.7%	0.2%	0.8%	0.5%	**96.2%**

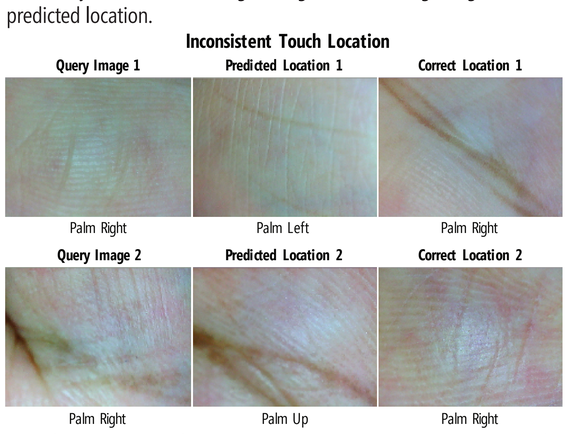
Classification percentages for classes at the fine-grained level (Stage 4 output), averaged across 20 folds and 30 participants. Each cell indicates the percentage of images assigned to a predicted class (column) for each actual class (row).

Examples of Classification Errors



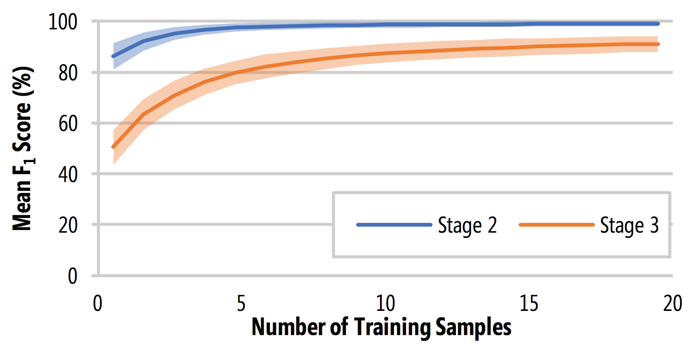
Classification errors were caused primarily by similarities between the locations' visual textures and poor image quality. Each set of images shows, in order, two examples (from different participants) of an incorrectly classified test image along with a training image from the predicted location.

Inconsistent Touch Location



Classification errors for several participants were also caused by inconsistent touch locations. Shown are two examples (from two different participants) where the locations were far enough apart to appear as entirely unrelated images.

Effect of Training Set Size on Performance



Although performance steadily improves as the number of training samples increases, our approach is able to function reliably even with only 5 samples per location—especially at the coarse-grained level. A less heuristic approach may be more elegant, but would likely require a significantly greater amount of training data.

Between-Person Classification

To look for similarities between participants, we conducted a between-person classification experiment. Our average F_1 scores were 71.7% at the coarse-grained level and 26.6% at the fine-grained level. Although these accuracies are clearly too low to support a reliable user interface without an individual training procedure, they may provide enough classification information to allow us to bootstrap the training set and reduce the amount of training data that is needed for a new user.

Between-person Coarse-grained Classification Confusion Matrix

	Palm	*Finger*	*Nail*	*Knuckle*	*Other*
Palm	**55.2%**	16.8%	7.8%	4.0%	38.8%
Finger	8.1%	**85.5%**	10.4%	2.1%	2.3%
Nail	0.2%	3.4%	**85.3%**	4.4%	0.9%
Knuckle	1.2%	0.2%	1.3%	**67.8%**	18.2%
Other	12.4%	4.1%	0.1%	18.2%	**60.3%**

MAKEABILITY LAB
inclusive design lab
HCIL Human Computer Interaction Laboratory
COMPUTER SCIENCE UNIVERSITY OF MARYLAND
UNIVERSITY OF MARYLAND

Lee Stearns University of Maryland, College Park, Computer Science
Uran Oh University of Maryland, College Park, Computer Science
Bridget J. Cheng Cornell University, Dept. Computer Science
Leah Findlater University of Maryland, College Park, Information Studies
David Ross Atlanta VA R&D Center for Vision & Neurocognitive Rehabilitation
Rama Chellappa University of Maryland, College Park, Electrical & Computer Engineering
Jon E. Froehlich University of Maryland, College Park, Computer Science
{lstearns, uranoh, leahkf, chella, jonf}@umd.edu, bjc267@cornell.edu, ross0128@bellsouth.net