frontiers in Robotics and AI

ORIGINAL RESEARCH
published: 20 November 2017
doi: 10.3389/frobt.2017.00061

PastVision+: Thermovisual Inference of Recent Medicine Intake by Detecting Heated Objects and Cooled Lips

Martin Cooney and Josef Bigun*

Intelligent Systems Laboratory, Halmstad University, Halmstad, Sweden

This article addresses the problem of how a robot can infer what a person has done recently, with a focus on checking oral medicine intake in dementia patients. We present PastVision+, an approach showing how thermovisual cues in objects and humans can be leveraged to infer recent unobserved human–object interactions. Our expectation is that this approach can provide enhanced speed and robustness compared to existing methods, because our approach can draw inferences from single images without needing to wait to observe ongoing actions and can deal with short-lasting occlusions; when combined, we expect a potential improvement in accuracy due to the extra information from knowing what a person has recently done. To evaluate our approach, we obtained some data in which an experimenter touched medicine packages and a glass of water to simulate intake of oral medicine, for a challenging scenario in which some touches were conducted in front of a warm background. Results were promising, with a detection accuracy of touched objects of 50% at the 15 s mark and 0% at the 60 s mark, and a detection accuracy of cooled lips of about 100 and 60% at the 15 s mark for cold and tepid water, respectively. Furthermore, we conducted a follow-up check for another challenging scenario in which some participants pretended to take medicine or otherwise touched a medicine package: accuracies of inferring object touches, mouth touches, and actions were 72.2, 80.3, and 58.3% initially, and 50.0, 81.7, and 50.0% at the 15 s mark, with a rate of 89.0% for person identification. The results suggested some areas in which further improvements would be possible, toward facilitating robot inference of human actions, in the context of medicine intake monitoring.

Keywords: thermovisual inference, touch detection, medicine intake, action recognition, monitoring, near past inference

OPEN ACCESS

Edited by:
Alberto Montebelli, University of Skövde, Sweden

Reviewed by:
Sam Neymotin, Brown University, United States
Per Backlund, University of Skövde, Sweden
Fernando Bevilacqua, University of Skövde, Sweden
(in collaboration with Per Backlund)

****Correspondence:***
Martin Cooney
martin.daniel.cooney@gmail.com

Specialty section:
This article was submitted to Computational Intelligence, a section of the journal Frontiers in Robotics and AI

Received: *15 May 2017*
Accepted: *02 November 2017*
Published: *20 November 2017*

Citation:
Cooney M and Bigun J (2017) PastVision+: Thermovisual Inference of Recent Medicine Intake by Detecting Heated Objects and Cooled Lips. Front. Robot. AI 4:61. doi: 10.3389/frobt.2017.00061

1. INTRODUCTION

This article addresses the problem of how a robot can detect what a person has touched recently, with a focus on checking oral medicine intake in dementia patients.

Detecting recent touches would be useful because touch is a typical component of many human–object interactions; moreover, knowing which objects have been touched allows inference into what actions have been conducted, which is an important requirement for robots to collaborate effectively with people (Vernon et al., 2016). For example, touches to a stove, door handle, or pill bottle can occur as a result of cooking, leaving one's house, or taking medicine, all of which could potentially be dangerous for a person with dementia, if they forget to turn off the heat, lose their way, or make a mistake. Here, we focus on the latter problem of medicine adherence—whose

importance has been described in the literature (Osterberg and Blaschke, 2005) and which can be problematic for dementia patients who might not remember to take medicine—and in particular on oral medication, which is a common administration route. Within this context, it is not always possible for a robot to observe people due to occlusions and other tasks the robot might be expected to do; thus, the capability to detect what a person has recently touched, from a few seconds to a few minutes ago, would be helpful.

Two kinds of touching typically take place during medicine intake: touches of a person's hands to packaging to extract medicine, and touches of a person's mouth to medicine and liquids to carry out swallowing. To detect such intake autonomously, electronic pill dispensors are being used, which feature advantages such as simplicity and robustness to occlusions, and are more accurate than administering questionnaires. A downside is that pill dispensors can only detect if medicine has been removed, and not if a person has actually imbibed; i.e., the first, but not the second kind of touching. In our previous work, we proposed a different option involving thermovisual inference: namely, that a system could combine knowledge of where people have touched by detecting heat traces, and what touches signify by detecting objects, to infer recent human–object interactions (Cooney and Bigun, 2017). In the current work, we propose how this approach can be extended by also detecting cooling of a person's lips, thereby detecting both kinds of touch occurring in medicine intake, as shown in **Figure 1**. A challenge was that it was unclear how an algorithm could be designed to detect such touches in a typical scenario in which both foreground and background can comprise regions of similar temperature, for several seconds after contact had occurred.

Based on this, the contribution of the current work is exploring the problem of detecting recent touches to objects and people in the context of oral medicine intake:

- We propose an approach for touch detection on objects and humans, PastVision+, which uses some simplified features in combining object and facial landmark detection, to handle scenarios in which touched objects can be in front of a warm background.
- We provide a quantitative evaluation of our approach for touch detection, in terms of how long touches can be detected on objects and people by our algorithm; moreover we also demonstrate performance of detecting some touches conducted by various people pretending to take medicine, while inferring actions and identifying individuals.
- We make freely available code and a small new dataset online at http://github.com/martincooney.

We believe that the resulting knowledge could be combined with existing approaches as a step toward enabling monitoring of medicine adherence by robots, thereby potentially contributing to the well-being of dementia patients.

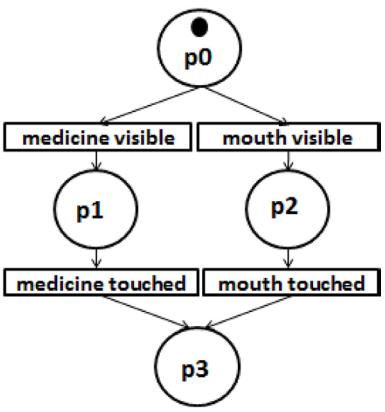

FIGURE 1 | A simplified Petri net process model describing thermovisual inference of oral medicine intake: A robot can move to view medicine packages and a person's lips, and infer that medicine intake might have taken place if both exhibit signs of recent touching.

2. MATERIALS AND METHODS

To detect touches on objects and humans, we proposed an approach depicted in **Figure 2**, implemented a rule-based version and a more automatic version of our algorithm, and conducted some exploratory tests.

2.1. Detecting Touches to Objects

The core concept of the proposed approach is described in the field of forensics by Locard's exchange principle: "whenever two objects come into contact with one another, there is always a transfer … methods of detection may not be sensitive enough to demonstrate this, or the decay rate may be so rapid that all evidence of transfer has vanished after a given time. Nonetheless, the transfer has taken place" (Ruffell and McKinley, 2005). In the current case, the property of interest being transferred is heat, which can be transferred through conduction, as well as through convection and radiation, in accordance with the second law of thermodynamics (Clausius, 1854): When a human touches cold objects, the objects become warmer, and the human becomes colder. Furthermore, after touching, heated objects cool in the surrounding air, returning to their original temperatures, whereas thermoregulation also works to restore human body temperatures. Throughout this time, the objects and human emit radiation as a function of their temperatures, which can be detected by sensors. Various equations can be used to gain some extra insight into this process. For example, flow of heat during conduction and convection depends on various factors such as contact pressure and surface properties, but can be approximately described *via* Newton's law of cooling. As shown in equation 1, this states that thermal energy Q transferred is proportionate to a coefficient h, the heat transfer surface area A, and the temperature difference, and can be used to predict, e.g., when a heated object will reach a specific temperature. Furthermore, Wien's displacement law in equation 2 indicates that objects at room temperature and people will mainly emit radiation in the long-wave infrared band which can be perceived by our camera, and can be used to check the predominant wavelength, λ_{max}, emitted by some object of interest at temperature T, where b is a constant.

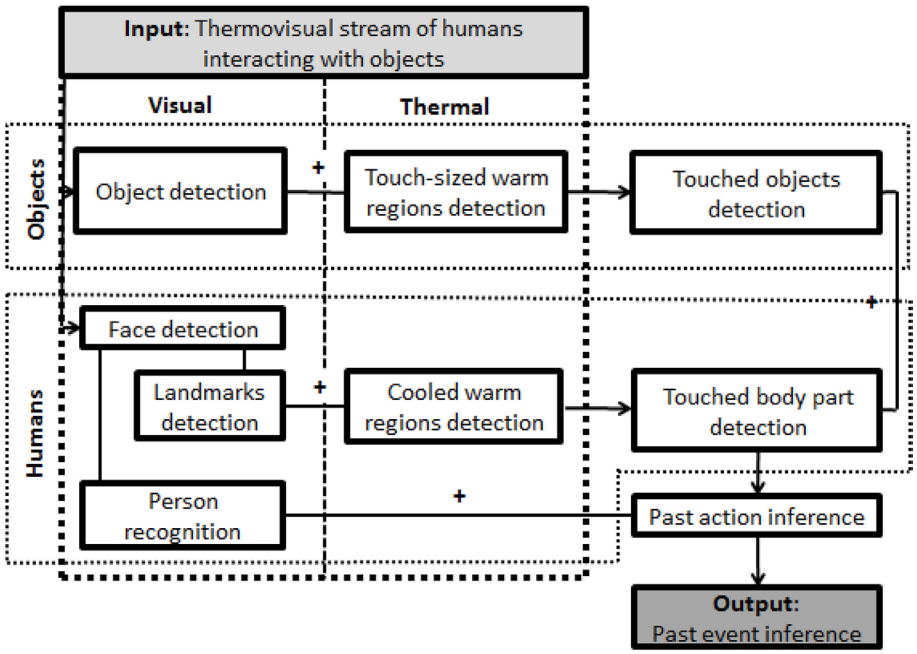

FIGURE 2 | PastVision+: process flow.

$$\frac{dQ}{dt} = hA\Delta T(t) \quad (1)$$

$$\lambda_{\max} = \frac{b}{T}. \quad (2)$$

Within this context, in previous work we investigated the first case above, of detecting heat traces on objects (Cooney and Bigun, 2017). A simplified context was assumed in which images could contain both humans and cool objects, but only in front of a cool background; based on this assumption, object detection was used to determine where in an image to look for heat traces, to ignore confounding heat sources such as nearby humans and thermal reflections.

In order to bring our approach a step closer to being used in the real world, we required a way to modify our approach so that this assumption would not be necessary. In particular, touch detection should also operate when an object is situated between the robot and a human. A challenge was that simply thresholding to remove image regions at human skin temperature did not work. This is because humans typically wear clothes, and the surface of clothed body parts can be around the same temperature as touched objects. As an initial solution to the challenge, we adopted the extension shown in **Figure 3**, of our previous algorithm (see examples in **Figures 4**–**5**):

1. Record an RGB and thermal image.
2. Register the images using a simple mapping determined ahead of time.
3. Detect objects within the RGB image; ignore objects which are not of interest for the application, such as dining tables, chairs, and sofas.

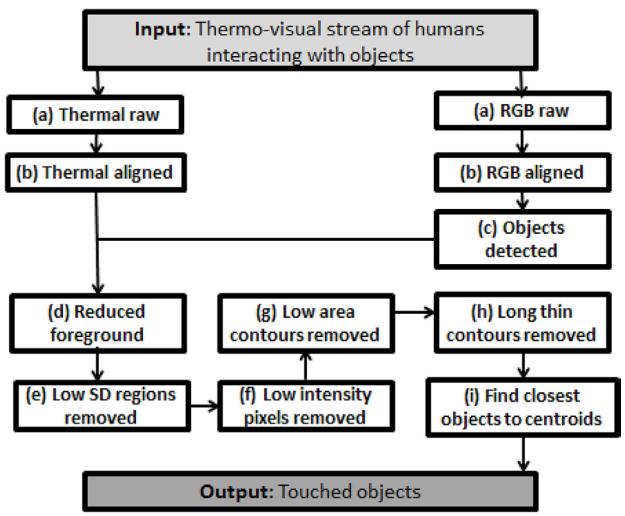

FIGURE 3 | Object touch detection process flow for the rule-based version of our approach. Note: the automated version of our approach combines steps (e)–(h).

4. For each region of interest containing an object, shrink the region slightly to avoid extra spaces.
5. Per foreground region, compare the standard deviation (SD) to a parameter θ_1, to ignore uniform areas unlikely to have been touched.
6. Per foreground region, threshold to extract pixels with intensity higher than the mean plus a parameter θ_2.
7. Per foreground region, find contours, to reject small noise with area less than θ_3.

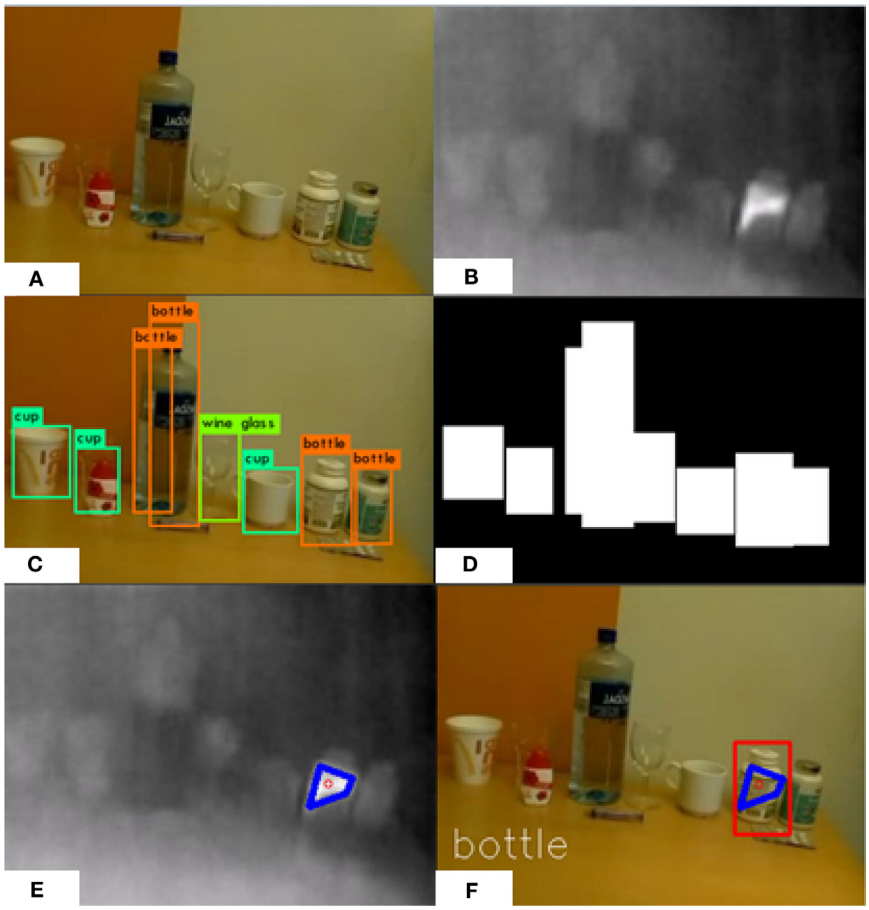

FIGURE 4 | An example of inference from single images in a simple case. **(A)** RGB image, **(B)** thermal image, **(C)** objects detected, with a few false positives and negatives, **(D)** initial mask image from bounding boxes to reduce noise, **(E)** thermal image with touched region and contour centroid detected, and **(F)** touched object identified.

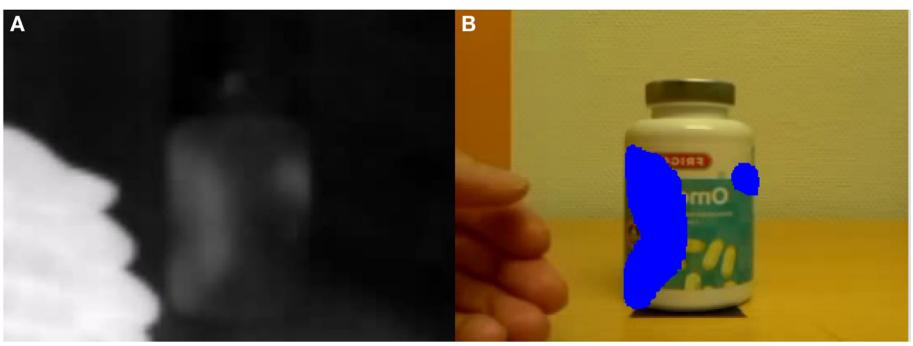

FIGURE 5 | A simple example of inference from a video, extracting heat traces and not humans, *via* thresholds, morphology, and a basic shape model for touching: **(A)** thermal image and **(B)** RGB image with heat traces drawn by the algorithm.

8. Per foreground region, calculate the arc length of the contour and surface to area ratio, to reject long thin contours with a surface to area ratio greater than θ_4.
9. Per foreground region, find the center of the touched region and set as the touched object the one with the least distance.

Registration in our case can be described by the linear transformations in equation 3, where (x', y') is a new aligned point derived from a point (x,y) in the original image, s_x and s_y are scaling parameters, t_x and t_y are translation parameters, and θ is a rotation parameter. Parameters were found by simply

viewing an overlay of the two image streams, visual and thermal, and pressing keys to alter parameters until images were aligned. Although more robust and complex approaches could be followed, such as compensating for intrinsic parameters such as lens distortion and calibrating with a thermal mask held in front of a heat source (Vidas et al., 2013), we expected our approach would be sufficient to consummate our goal of obtaining some basic insight regarding the feasibility of thermovisual inference of medicine intake.

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} s_x cos\theta & -sin\theta & t_x \\ sin\theta & s_y cos\theta & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{3}$$

The bounding boxes found as a result of object detection are used to exclude noise arising from thermal reflections and human heat. However, many objects such as medicine bottles are not rectangular and cannot be cleanly segmented *via* bounding boxes. For this, the bounding boxes are shrunk slightly. Also, the SD is checked to discard thermally uniform regions based on an expectation that touched objects will comprise touched and untouched regions differing in thermal intensity. The mean is used to find relatively warm areas within a region of interest based on the assumption that touched areas are typically warmer than untouched areas; using a mean to threshold instead of a fixed parameter is useful because heat traces cool over time and some objects can be warmer than others, e.g., due to warming by the sun or heat dissipated by electronic devices. Area and surface area are used to discard noisy small contours and long thin contours which can arise from the outlines of objects located in front of warm backgrounds, based on an assumption that touches with fingers and hands will tend to be a certain size and shape.[1]

For object detection, object classes and locations were calculated simultaneously using a single convolutional neural network with many layers trained on a mixture of detection and recognition data, in conjunction with some non-maximum suppression (Redmon et al., 2016). The architecture consisted of 24 convolution layers and 2 fully connected layers, with leaky rectified linear activation in all but the final layer, and output was optimized in regard to the sum of squared error.

Thus, we proposed an approach to deal with scenarios in which both humans and objects are visible to the system, using some common features in conjunction with object detection.

2.2. Detecting Touches to People

Touches to objects might not always be sufficient to infer what has been done. For example, a person with dementia could extract medicine from a package and forget to take it, or accidentally take someone else's medicine. In the current article, we propose that additional insight can be drawn from also detecting the after-effects of touching on a human's body. For example, cold or warm lips could indicate liquid intake, which is common for oral medicines such as pills and syrups, cold skin could result from applying cream, cold around the eyes could come from applying eye drops, a warm back could indicate sitting or reclining, and cold hands could indicate that some object-related action has been performed. Here we focus on a limited scenario involving just one indicator, cooling of the lips, as a start for exploration; also we note that a robot can move to seek to detect touches on objects or humans if they are not visible, but this is outside of the scope of the current article.

To detect lip cooling, our approach involved detecting facial landmarks and then computing some features from within regions of interest to find anomalies. To detect facial landmarks, faces were first detected by considering local gradient distributions over different scales and parts of an RGB image, from which regression functions were used to iteratively refine location estimates (King, 2009; Kazemi and Sullivan, 2014). Specifically, Histogram of Oriented Gradients (HOG) features were derived by finding gradient magnitude and orientation information for each pixel, g_{pi} as in equation 4, forming histograms within image cells, normalizing over larger blocks, and feeding the descriptors to a Support Vector Machine classifier with decision function f(x), learned parameters α_i and b, and a linear kernel K as in equation 5, where x_i and y_i are training data and labels. Then, a cascade of regressors was used to iteratively improve initial estimates of landmark locations as in equation 6, where r_i is the ith regressor, I is the image, and L_i is an estimate of the landmark locations based on regressors r_{i-1} to r_0. Regressors were learned *via* gradient boosting by iteratively combining regression trees, as piecewise constant approximating functions, with splits greedily chosen from randomly generated candidates to minimize the sum of squared error.

$$g_{p_i} = \left(\sqrt{g_x^2 + g_y^2}, arctan\left(\frac{g_y}{g_x} \right) \right), \quad where \quad g_x = \frac{\partial I}{\partial x}, g_y = \frac{\partial I}{\partial y} \tag{4}$$

$$f(x) = sgn\left(\sum y_i \alpha_i K(x_i, x) + b \right) \tag{5}$$

$$L_{i+1} = L_i + r_i(I, L_i). \tag{6}$$

Some minor problems arose with facial landmark detection. For example, there was a delay in recording thermal and RGB images, which meant that sometimes when a person was moving quickly, faces detected in the RGB image did not match the facial region in the thermal image. To deal with this we implemented an Intersection over Union (IoU)-like metric to check that alignment was acceptable: a threshold on thermal pixel intensity was used in the vicinity of a detected face and the thermally found contour compared with a contour found from the landmarks detected in the RGB image, under the assumption that there would be nothing else at the temperature of bare human skin directly behind a person's face. Another problem was that sometimes a person's chin was detected as a mouth. To deal with this we implemented a simple check on the length of the face using a threshold parameter. And, in some frames faces were not detected. For this we implemented some temporal smoothing, combining processing results for three frames instead of just one.

[1] We also tried backprojection based on extracting a histogram from near the center of the bounding boxes but encountered difficulty with some objects such as packages which contained many different colors, transparent glasses, and reflective cups. Other approaches might be to model the shapes of certain objects, consider symmetry, or use snakes.

From the facial landmarks, we computed some simple statistics based on the intensities of the pixels within the lip region. In pretests we noticed problems when incorporating the inside of the mouth, which can be hot near folded tissue like the base of the tongue or cold due to saliva. Because people often open their mouths to talk, this area inside the lips was excluded.

In addition to touch action detection, cropped face rectangles were also classified to determine the identity of the person taking medicine. Local Binary Pattern Histograms were used as features, for robustness to global monotonic lighting changes. Binary codes were extracted by comparing the intensity of each pixel to that of its neighbors and then forming histograms within facial regions, which were again passed to a classifier as described in equation 5.

2.3. Implementation Tools: Software and Hardware

To implement inference, various software and hardware tools were used, as summarized in **Table 1**.

The Pylepton library[2] was used to access information from our thermal camera. OpenCV[3] was used for general functions such as thresholding and finding contours. YOLO was used for object detection due to ease of use and speed (Redmon et al., 2016). The dlib library[4] was used with OpenCV trained on the 300-W face landmark dataset (Sagonas et al., 2013) to detect facial landmarks, likewise due to ease of use and speed (King, 2009). OpenCV (code by Philipp Wagner[5]) was used for face recognition to identify persons for simplicity, and data were prepared in the style of the AT&T Database of Faces, as small grayscale images in PGM format. Scikit-learn was used for classification in the automated version of our approach (Pedregosa et al., 2011).

For hardware, we used an off-the-shelf inexpensive thermal camera and RGB camera attached to a Raspberry Pi 3, and a remote desktop for processing. The thermal camera had a resolution of 80 × 60 which we felt was sufficient for our purpose and was designed to detect temperatures typically present in human environments and human bodies (8–14 μm). Some unoptimized code we wrote showing various data streams while recording thermal, RGB, and time data ran at approximately 8.6 fps. Processing was conducted on a desktop with an i5 2400 CPU @ 3.1 GHZ.

[2]https://github.com/groupgets/pylepton.
[3]http://opencv.org.
[4]http://dlib.net/.
[5]http://www.bytefish.de/dev/libfacerec/.

TABLE 1 | Tools.

Component	Tool
Thermal data acquisition	Pylepton
General image processing	OpenCV
Object detection	Darknet/YOLO
Facial landmark detection	Dlib
Pattern recognition	Scikit-learn
Thermal-visual sensing	FLIR 80 × 60, 8–14 μm; standard small RGB camera

Thus, although our approach is not dependent on any of these tools—there are many options available for each component module—we found that they worked for our application.

Additionally, we implemented two versions of our approach, the first rule-based, and the second more automatic. The rule-based version is easier for humans to interpret and can be structured more easily, e.g., to have a low rate of false positives. The automatic version requires less work to find parameters.

2.4. Evaluation 1: Touch Detection

The system we had implemented needed to be evaluated to clarify basic feasibility. Touch detection, the heart of the current system, was selected as the focus of evaluation: How long could traces be detected on objects and people by the algorithm or by a human observer, and how often will the system correctly/incorrectly guess that an object/person was touched? Also, for detecting touches on people we were curious to see which features would perform well. To answer these questions, we first conducted a basic check, observing how long touches could be seen in the thermal feed with the naked eye on both objects and body parts as shown in **Figures 6–8**. The basic check was promising so we proceeded to check the performance of our rule-based algorithm for (a) objects and (b) people.

(a) An experimenter sitting on a sofa in front of a table with some objects touched the objects one at a time, and the algorithm was invoked to detect touches. Thermal and RGB data, as well as times for each frame tuple, were recorded. Objects were touched one at a time, for approximately 5 s, with a 2-min waiting period afterward for each. Five objects were touched four times each: a labeled PET pill bottle, a labeled HDPE lotion bottle, a paper box, a glass cup, and a ceramic cup. Afterward, the recorded videos were used to manually identify frames just before and after touches. Then frames were extracted automatically at 5, 10, 15, 30, 60, and 90 s after touches ended. Furthermore, as described in the previous section, the system chose three frames for each time period for robustness. This resulted in 480 images extracted for analysis.

Objects were touched one at a time to simplify evaluation. Touches were made to last 5 s as we expected people would touch medicine packages and cups this long. Two minutes was chosen based on a pretest indicating touches were hardly visible after this time. The longest time period, 90 s, was chosen to fall within the 2-min period. The objects were chosen to each represent a different material and to cover some typical medicine packages relating to oral intake and topical application. We chose a challenging scenario in which there can be similar thermal patterns in both foreground and background because we felt this is a fundamental scenario which our algorithm should be able to handle.

(b) To evaluate system performance in recognizing touches on people, the experimenter drank water, and the algorithm was invoked to detect touches to the experimenter's mouth. Two cases were investigated for the water's temperature:

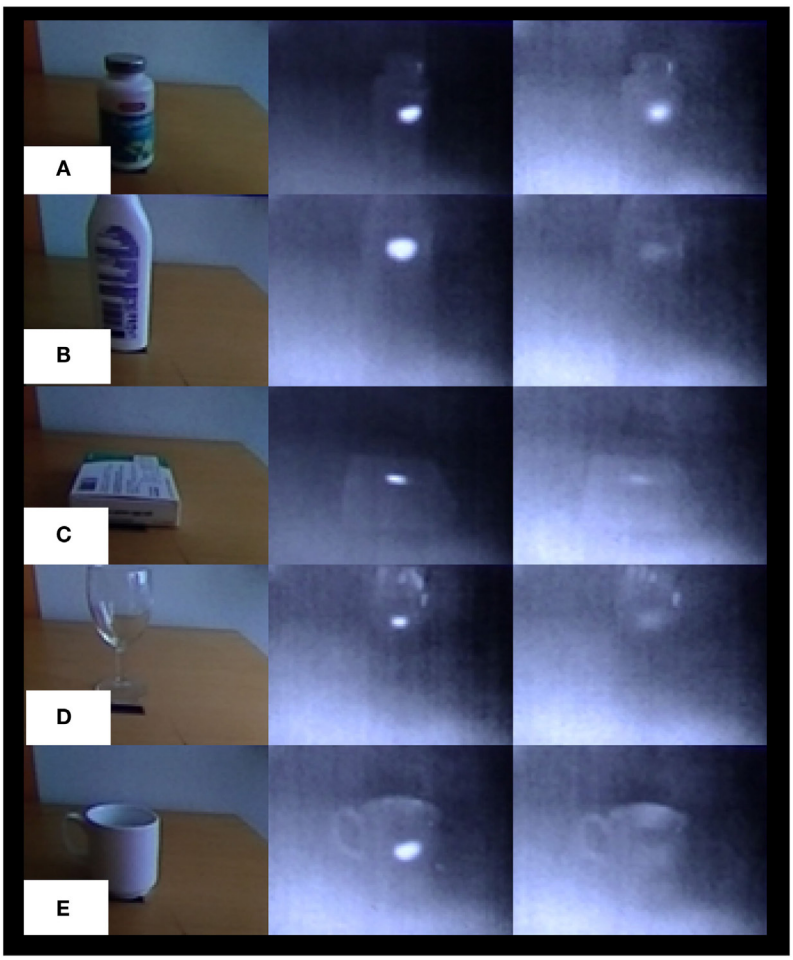

FIGURE 6 | Heat trace decay over time for some typical materials for medicine intake: for **(A)** PET, **(B)** HD-PE, **(C)** paper, **(D)** glass, and **(E)** ceramic; RGB data (left), thermal data soon after touching (center), and thermal data after 30 s (right).

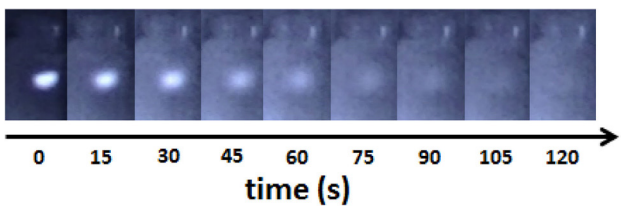

FIGURE 7 | Heat trace decay over time for PET. After 2 min the heat trace was difficult to see.

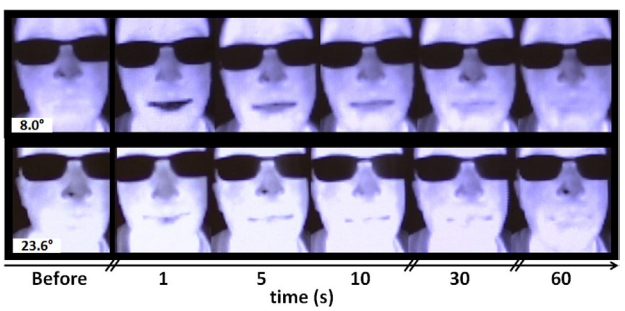

FIGURE 8 | Cooling lips after drinking cool and room-temperature water. Consent has been obtained from the individual depicted (the first author) for the publication of this image.

cold (refrigerated), and room temperature. For each case, the experimenter drank five times. Afterward, the recorded videos were used to manually identify frames just before and after touches. Then frames were extracted automatically at 5, 10, 15, and 30 s after touches ended. Furthermore, as described in the previous section for each time, the system chose three frames for robustness. This resulted in 180 ($5 \times 6 \times 3 \times 2$) images extracted for analysis. Some features were also checked along the way: means, medians, SDs, and "deltas" were calculated based on the intensities of the thermal-image pixels within the lip area, where deltas compared the mean thermal intensity within the lip area to that of the area from the bottom lip to the chin.

The longest time of 30 s was chosen based on our expectation that after 30 s cooling is not easily detected, which came from a pretest we conducted, and also because the area inside the lips is quite small and thus noise effects could be quite strong when there is little contrast. The two temperatures of water were chosen because we expected that this would be common for oral intake; we also wanted to know how different the results would be and if detection is feasible only with cold water, or also with tepid water. We chose a challenging scenario in which both objects and human are seen and the sensor is not close because we feel this is a fundamental scenario which our algorithm should be able to handle. Also, we selected some typical features to check because we expected them to perform well: Means were expected to capture when touches occur without knowing exactly where or which pixels will change in intensity, because we know that heat flow will occur during medicine intake in some portion of objects and a person's lips due to the difference between body and room temperature. SDs were expected to be low in the absence of touches due to thermal diffusion within objects and a person's lips, and high immediately after touches due to the difference in temperatures in touched and untouched regions. Medians are less susceptible to outliers than means. Deltas were selected because in pretests we sometimes observed some global monotonic temperature changes, which absolute means and medians would not be robust to; these might have been due to breezes or cooling from the person waiting without moving.

The lip region and not the entire mouth region was selected because we observed noise in preliminary tests when people opened their mouths to talk, as areas where tissues touch such as under the tongue can be warm, whereas other areas can be wet and cool.

For deltas, the lip region was compared to a region between the lips and chin because the latter region seemed stable in pretests, and other options seemed problematic: the nose region is tricky due to breathing, eye temperatures cannot be detected for people wearing glasses, the forehead region can be occluded by hair, with much variance in people's haircuts, and left or right cheeks vanish when a face is turned sideways and have large borders, where inconsistencies can occur. **Figure 9** shows the data acquisition setup and the system detecting drinking.

2.5. Evaluation 2: General Feasibility

The experiment described in the previous section was intended to provide some exploratory insight into a core question for thermovisual inference of medicine intake by assessing how long touches could be detected on different materials and at some different temperatures; but, the general significance of the results and performance of the other parts of the algorithm were unclear due to the simplified context. For example, touches to objects had involved merely pressing a hand against the front of a container, which might be different from touches arising from medicine intake. Also, touches had only been performed by a single person, which might be different from touches performed by other people. And, some manually found parameters were used in the rule-based version of our algorithm, but parameters might change in different contexts. Moreover, the feasibility of discriminating actual medicine intake from similar confounding touches, and of identifying individuals taking medicine, had not been addressed.

Thus, to gain some further insight into the general significance and feasibility of the proposed approach, an additional simplified check was conducted. To investigate if touches arising from medicine intake can be detected, ten participants at our university (five females and five males; average age = 31.6 years, SD = 5.7 years) were asked to pretend to take medicine while seated in front of our camera; touches were then detected by the automatic version of our algorithm. To assess the other parts of our approach, participants were also asked to absent-mindedly touch a medicine package, which might be difficult to distinguish,

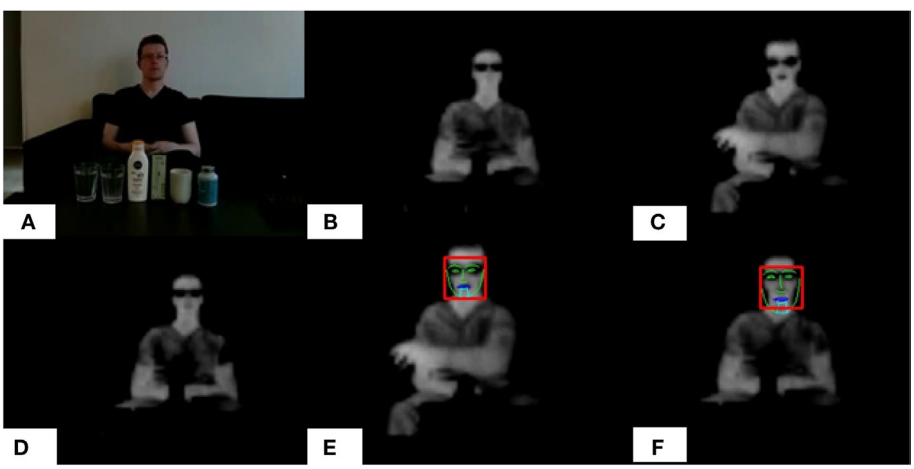

FIGURE 9 | Detecting drinking. **(A)** RGB image of setup with a human and some objects. **(B)** Thermal image before drinking **(C)** Thermal image after drinking, **(D)** Thermal image 30 s after drinking, **(E)** processed thermal image, indicating face rectangle (red), face outline (green), lips area (blue), and area below the lips (cyan), and **(F)** a model for normal face length was used to avoid some failed detections which detected a mouth at the person's chin. Consent has been obtained from the individual depicted (the first author) for the publication of this image.

and to sit together in pairs, to check if the algorithm could correctly identify who had pretended to take medicine.

At the start of sessions, each pair of participants was seated at a table in front of our camera, a medicine package, and two cups filled with water, as shown in **Figure 10**. After reading a short description of the data acquisition session indicating that the task would be to pretend to take medicine and touch a container absent-mindedly, participants provided written consent; participants were not given instructions on how to take medicine or how to otherwise touch, and did not receive monetary compensation for participating. Next, the experimenter started recording data, and also filmed a monitor which showed the output of the thermovisual feed next to a clock, to be able to calculate how long participants took to perform actions. Then the experimenter read aloud instructions in a pregenerated random order which was different for each pair of participants. Thus, the independent variables manipulated were the type of action (medicine intake or absent-minded touch), and participant (A or B), which were grouped to form four cases per pair, two per participant, as shown in **Table 2**, for five pairs of participants. Participants then conducted actions using containers for medicine and water with some typical shapes and materials: a plastic medicine bottle and two glass tumblers. Two cups were used for hygienic reasons, so that each participant could drink from their own. After each action, a 2-min wait was scheduled, to allow temperatures to return to their previous levels and to be able to compare accuracies at different times after an action. During this time participants were allowed to talk and move during the session. In total, each session took approximately 15 min.

After the sessions, frames for 18 actions were extracted as in the previous experiment—before actions, immediately after, and at 5, 10, 15, 30, and 60 s after—yielding seven frames per action, and 126 frames in total; for person identification, face detection was applied to these frames to automatically extract 245 face images. Then the automatic version of our algorithm was applied to the extracted frames. Features were chosen based on those reported in the last section for the rule-based version of our system: For detecting touches to objects, we used the mean, SD, median, minimum, and maximum intensities, as well as the area of the maximum detected contour and minimum surface to area ratio, within the segmented region of the thermal image corresponding to the medicine package. And for detecting touches to lips, we used the mean, SD, median, minimum, and maximum intensity of the lip region, and the mean and SD of the intensities of the thermal image in the region below the lips. For person identification, along the way we also explored the possibility of using some simple thermal features to detect characteristics such as glasses, haircuts and beards, and a person's size, which could facilitate recognition: the mean and SD of thermal intensities in the eye regions which might be blocked by glasses or hair, the area of the thermal contour at the top of the face which might describe the person's haircut, the mean of thermal intensity for the area below the lips where a beard would be, and the area and height of the thermal face contour. The features were then input to a classifier which was used to conduct leave-one-out crossvalidation to compute accuracies, defined as the number of correct predictions divided by the number of samples. For this, k-nearest neighbors (k-NN) was used as a simple check, as it is easy to use and provides consistent results, with an error rate not worse than twice the minimal Bayes error rate as the number of training data approaches infinity (Cover and Hart, 1967). This algorithm finds the distance from a test case to the training data, and associates the label of the closest found training data to the test case. For action inference, rules were used: If an object touch was not detected, the system output that an action had not occurred. If an object touch was detected but no lip touch was detected, the system output was set to absent-minded touch. If both kinds of touches were detected, the system output that medicine intake had occurred.

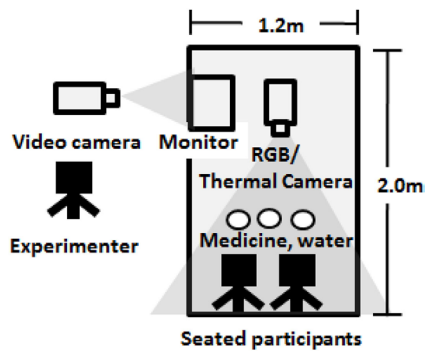

FIGURE 10 | Experiment setup for second evaluation.

TABLE 2 | Cases per pair of participants (A and B).

Person	Touch type
A	Pretended to take medicine from a bottle and drank water from a cup
A	Absent-mindedly touched the bottle
B	Pretended to take medicine from a bottle and drank water from a cup
B	Absent-mindedly touched the bottle

3. RESULTS

3.1. Results 1: Touch Detection

(a) Accuracies for each material over time can be seen in **Figure 11**, and averages in **Table 3**. At 15 s, half or less of touches were detected. At 60 s the touch detection rate was approximately 0%, with a slightly higher rate of 5% for the glass cup which we expect might have been due to noise. Thus, touches to objects appeared to be best detected within a short time of 0–10 s. However, we note that our scenario was highly challenging; if there is no human in the background we expect better accuracies and longer detection times. Also, as expected, touches no longer detected by the algorithm could still be seen with the naked eye, suggesting that better performance can be expected in the future; the problem was segmenting the objects from the warm background. For materials, all the materials tested allowed some detection of touches. The PET pill bottle appeared to allow easiest detection of touches, and the ceramic cup was most difficult.

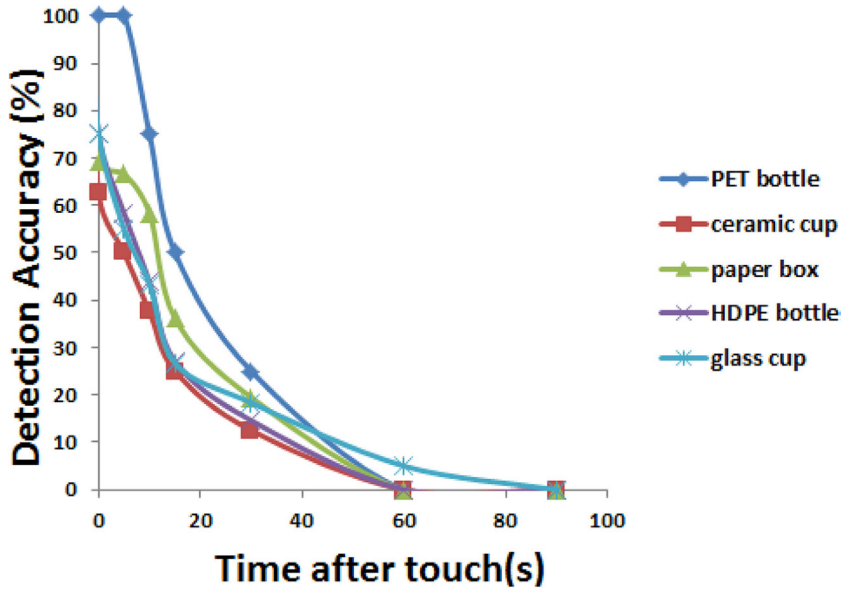

FIGURE 11 | Accuracy of detecting touches to objects, by material and time.

TABLE 3 | Detection accuracies of heat traces on objects over time.

Time (s)	Accuracy (%)
0	76.4 ± 14.2
5	66.0 ± 20.0
10	51.6 ± 15.2
15	33.0 ± 10.5
30	18.0 ± 4.8
60	1.0 ± 2.2
90	0.0 ± 0.0

Some examples of correctly and incorrectly identified cases are shown in **Figure 12**. The result was not perfect due to the difficulty of the problem: the additional processing required to exclude warm background areas also removed true positives at longer time periods.

(b) Results shed light on which features might be useful to detect touches on human body parts, and on the system's accuracy in detecting touches due to drinking liquids of different temperatures.

Figure 13 shows the performance of some features we compared. Delays seemed to exist in the responses for the mean and median for cold water, and the mean, median, and deltas for tepid water. This is possibly because the physical activity of drinking or closer distance to the camera when leaning forward to drink also affected detected temperatures. Also, there appeared to be little difference between means and medians, suggesting that drinking does not generate a small number of cold outlier pixels, but rather many cool pixels emerge. Thus, all the features seemed to be potentially useful, with the usefulness of SD features in particular suggested, which appeared to capture the changes in temperature quickly.

System accuracy, in detecting drinking cold versus tepid water, is depicted in **Figure 14**. In general, the results appeared to be in line with our expectation: the algorithm had less trouble discriminating touches caused by drinking cold water (91.7 versus 83.3%). The results were not perfect due to the difficulty of accurately detecting of the lip region. At 30 s the rate for tepid water seemed slightly better, but we expect this might have been due to noise. We also note that to compute the accuracy of our approach, frames for which a correct face was not detected were not considered. Factoring in such frames results in an unintuitive result where accuracy is lower in the cold water case (73.3 versus 83.3%) because faces were not detected in some frames for the cold water, but this was unrelated to the coldness of the water, as the thermal data is not used for detecting faces. Also in both cases drinking was sometimes detected at 30 s, which was longer than we expected, possibly because it was difficult to judge visually based on the small area of the mouth.

Thus, we confirmed our algorithm could detect heat traces on objects and people over several seconds in settings comprising both objects and people, for some typical object materials, in a simplified context. The data used for evaluation has been made freely available online with code for others, toward expediting progress in this promising area of near past inference.

3.2. Results 2: General Feasibility

The actions participants performed were complex and varied greatly. Medicine intake involved multiple touches with one or both hands, as shown in **Figure 15**, during which objects were moved around the table, and participants moved their lips while talking, smiling, or turning their heads. Eight of ten participants first extracted medicine then drank, but one filled her mouth with water before extracting medicine, and another drank while still holding the medicine bottle; some participants also smacked

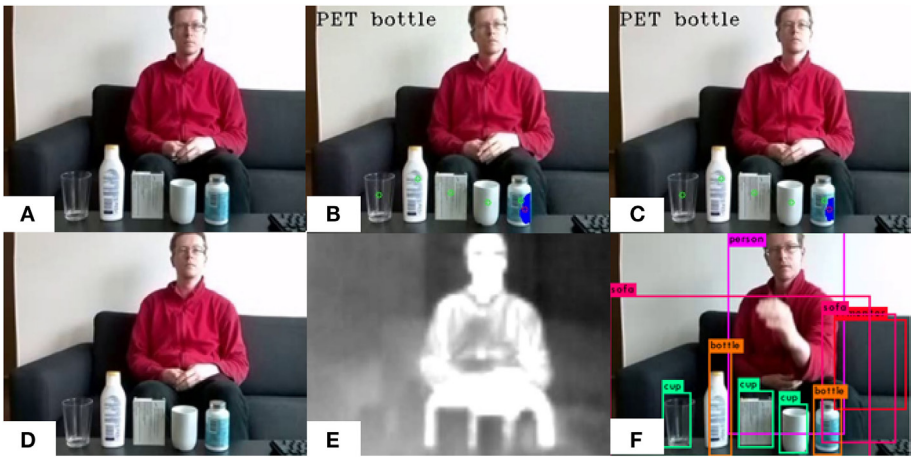

FIGURE 12 | Some examples of detecting touches to medicine packages. **(A)** Time −5 s: a true negative before touching, **(B)** time 0 s: a true positive immediately after touching (touch shown in blue by the algorithm), **(C)** time 30 s: the same touch after 30 s, where a decrease in area of the touch can be seen, **(D)** time 60 s: a false negative, the same touch is no longer detected, **(E)** time 60 s: the touch can still be seen by the naked eye in the thermal image, and **(F)** different touch: a rare false positive, caused by the black sofa in the background being mistaken as a TV monitor, and part of the experimenter's leg as a touch. Consent has been obtained from the individual depicted (the first author) for the publication of this image.

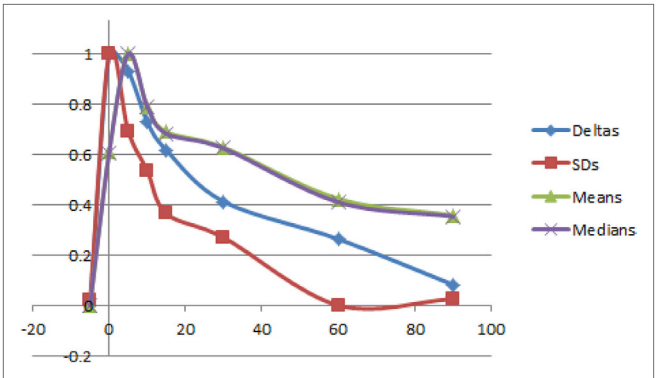

FIGURE 13 | Comparison of features over time while drinking water.

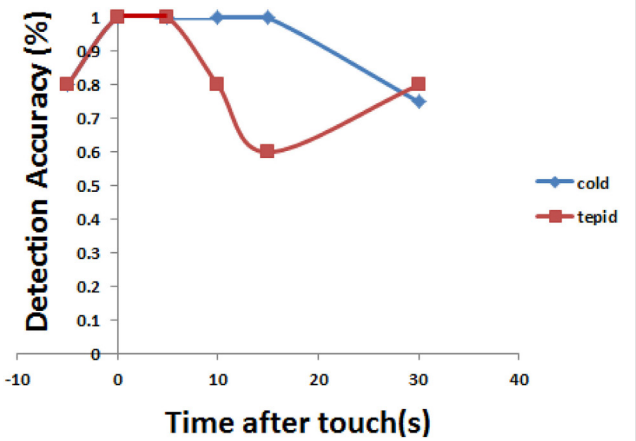

FIGURE 14 | Comparison of accuracies for cold and tepid water.

their lips after drinking. Likewise, absent-minded touches were sometimes complex, for example in picking up an object and rotating it in various directions, and ranged from quick squeezes to longer touches.

Average touch durations were also calculated using the video footage. Touches to the medicine package took 8.0 ± 3.0 s (maximum 13 s, minimum 4 s), touches to cups took a similar amount of time, 7.9 ± 3.1 s (max. 15 s, min. 4 s), and absent-minded touches were typically shorter, 4.8 ± 2.6 s (max. 8 s, min. 1 s). Thus, medicine intake took about 15 s, suggesting that our previous estimate that people would touch for about 5 s to take medicine had been conservative, also because we estimate that elderly persons with dementia might require more time than our participants. Absent-minded touching was faster than the two touches for medicine intake, possibly because typical interpretations such as moving the bottle a short distance or visually inspecting could be accomplished quickly.

Performance of our approach on the data is shown in **Figure 16**. Lip touch detection accuracy was 80.3% initially, in part because lips were usually visible to the camera; a slight increase in accuracy after some seconds could be partly due to participants still moving, noise, or a cooling effect as the lips dried and water evaporated. Object touch detection accuracy was lower, at 72.2% initially, decreasing to the random chance level at 15 s. We think there were several reasons for this: First, for some touches the heat trace was not clearly visible from the front, possibly due to the variance in participants' actions with respect to bottle placement, touch duration, temperature, and pressure. Second, our simplified approach for registration and segmentation did not always completely remove areas of warm background which could be mistaken for touches. Third, at the start for measuring, an average of 8 s had already passed since touching the bottle because time was taken for drinking water. Action detection was 58.3% initially, and remained higher than random chance (33%) until the 60 s mark. For

FIGURE 15 | An example of the complexity of medicine intake. The participant raises his left arm from his lap to allow him to reach forward while picking up the bottle with his right hand, opens the bottle with a spherical grip using the left hand, extracts a pill onto his left hand, replaces the bottle on the table and takes the medicine with the right hand to the mouth, reaches out for the water with his left hand while pushing down on the table with his right wrist, replaces the bottle and returns both hands to his lap. As a result of this action, the medicine bottle is in a new location and not fully closed.

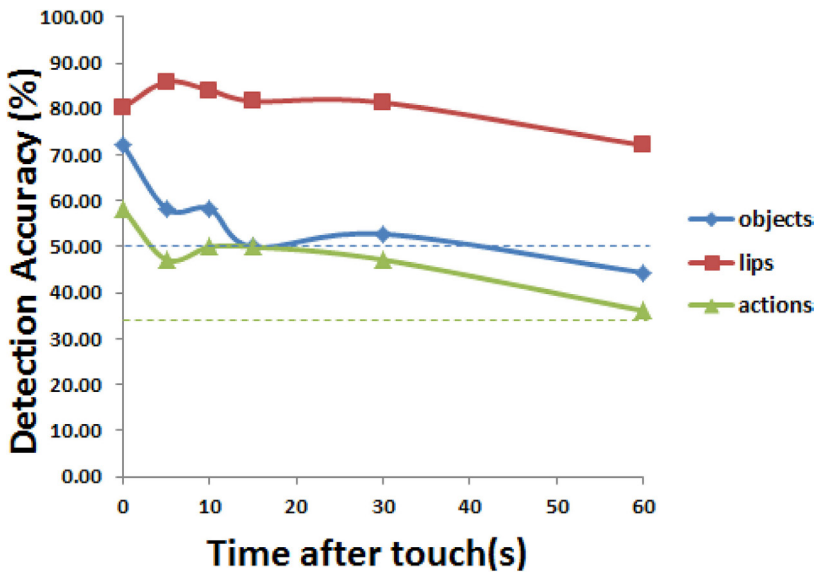

FIGURE 16 | Accuracies for touch detection on objects and lips, as well as action inference, over time.

person identification in our scenario, the thermal features yielded an accuracy of 89.0%, which was less than the off-the-shelf camera solution using only the RGB features and linear binary patterns (95.5%), possibly because illumination was not a problem; nevertheless the usefulness of thermal features has been described for some situations, e.g., involving low illumination or adversarial situations to identify makeup, masks, or computerized images.

We feel that these exploratory results, while highlighting some of the difficult challenges of inference of recent medicine intake, also suggest the promise of adding inferential capability to current systems, which do not detect lip touches and have no capability to detect touches conducted in the near past.

4. DISCUSSION

The current article described an approach, PastVision+, which uses thermovisual information from objects and humans to infer recent unobserved human–object interactions, in the context of medicine intake. PastVision+ allows quick inference from single images, which can be processed in several seconds, and requires only inexpensive components: a small thermal camera, costing approximately two hundred US dollars, a regular camera, and a computer. In our evaluation, we observed that our approach could be used in a challenging scenario in which some touches occurred in regions where the background was also warm, with about 50% detection accuracy of touched objects and 100% detection accuracy of cooled lips 15 s after a touch. Moreover, touches to typical objects used in medicine intake were visible in the thermal data for longer than 30 s, suggesting that improvement will be possible in the future. We also observed that some people's touches seemed to be easier to detect than others: an automatic version of our algorithm could detect touches conducted by various people pretending to take medicine at 72.2% accuracy for objects, 80.3% for lips, and 58.3% for action inference just afterwards, with a rate of 89.0% for person identification.

Performance was often not perfect due to the highly challenging scenarios and the exploratory nature of the current work. To compensate, we think such a system can be combined with existing methods like pill dispensors to improve the robustness of object touch detection, which seems to be important for improving accuracy; second, inference could be used, not just for humans to make decisions, but also by a robot to remind or investigate further, for which perfect accuracy might not be required. For combination with pill dispensors, we advise the use of plastic rather than metal dispensors, which allow longer touch detection durations; also individuals to be monitored could be provided with plastic cups rather than metal, or ceramic, cups, with shapes conducive to detection.

4.1. Related Work

Our work builds on various research that has been conducted on healthcare robots, past inference and thermal touch sensing, and intention inference.

Robot and intelligent assistive devices have been built to perform useful healthcare tasks like helping patients to move, fetch objects, or monitor people's health. Robots can pick patients up (Mukai et al., 2010), support their movements (Kawamoto et al., 2010), bring medicine and offer reminders (Pollack et al., 2002; Graf et al., 2009; Dragone et al., 2015), and monitor (Noury, 2005; Takacs and Hanak, 2008). In our own previous work we developed a robot which used ambient sensors to monitor falls and then moved to a fallen person to ask if they were okay (Lundstrom et al., 2015); another robot was built to detect and visualize fallen persons on a map also employing thermal means, as a pre-step for conducting first aid (Hotze, 2016). However, these robots cannot infer people's intentions.

One robot could infer a person's intentions to press buttons by observing the person's motions visually and adapting a model of itself to the person (Gray et al., 2005). Another robot was able to infer the intentions of other robots in a game of hide-and-go-seek, by correlating their motions (Crick and Scassellati, 2008). In our previous work we explored how to infer underlying motivations for social behavior directed to a small robot in interactions (Cooney et al., 2015). These systems were designed to infer based on observing current actions, and cannot infer past actions.

Past inference is common in many fields, such as astronomy, geology, archeology, and forensics, but has received little attention from the perspective of monitoring and robotics. One modality which allows past inference at time scales useful to humans in homes is thermal sensing, which has been explored from a different perspective as a user interface. For example, Benko et al. described a simplified strategy for how touches to a spherical display could be detected; normalizing was conducted to compensate for different intensities at different heights along the sphere, followed by thresholding and finding contours (Benko et al., 2008). Furthermore, some properties of heat traces on different materials—glass, tile, MDF, aluminum—were investigated by Abdelrahman et al. (2015). The paper most similar to the current article could be argued to be one by Larson and colleagues, which proposed various mechanisms and explored various properties of thermal sensing, also for a user interface application (Larson et al., 2011). In their approach, HeatWave, a person's hands were detected to reduce the search space for heat traces; per-pixel temperatures, temperature changes over time, and temperature differences to a background were used to identify heat traces. Additionally the authors explored multiple finger touches, hovering, shapes, pressures, and reflections. This resulted in much new knowledge, but this knowledge did not indicate how to conduct past inference: e.g., to detect touches to objects in front of warm backgrounds from single images, or to detect touches to specific parts of people's bodies.

We also noted some alignment between our results and theirs, in that touches to objects appeared to be best detected within a short time of several seconds, although reasons differed. In their study, it was assumed that an appropriate surface would be available to write on; therefore, challenges other than segmenting foreground touches from warm backgrounds were addressed, such as detecting quick or light touches. In our case, such touches were not expected because opening packages and drinking requires several seconds; therefore, we addressed the different challenge of warm backgrounds which was important for our context. Likewise, different features were used in our studies, but it is unclear if duration of touch detection would be lengthened by changing these, as selection was motivated by the context; for example, computing contour length to area ratios might not be useful for drawing heat traces with a finger, if there are no artifacts in front of a warm background generating long thin noise.

However, if we had considered a simple background or Larson et al. had considered touches lasting several seconds, we expect touches would have been detected longer, also because we could see the result of touching for longer times in the thermal image during a pretest we conducted. Regardless, we feel that, given a typical robot locomotion speed of 1 m/s which is slightly slower than human walking speed, several seconds could be enough in some cases for a robot to move to avoid an occlusion and infer what a person has done.

4.2. Limitations

The results of this feasibility investigation are limited by the simplified scenario, and it should not be interpreted that our approach is ready to be applied in the wild. Human environments typically contain various warm objects, from computers and cell phones to pets; temperatures constantly fluctuate; and heat reflections and heat contamination of frequently touched surfaces complicate thermal touch sensing. Likewise, illumination and occlusions are problems for RGB cameras; object detection against complex backgrounds, facial landmark detection for extreme or occluded poses, and person identification are also highly challenging. Objects can occlude one another, such as in cluttered medicine cabinets, flat objects can be hard to see, and touches to metals or warm objects are hard to detect. People also constantly move, conducting adapter gestures like scratching themselves or licking lips, exercising, opening or closing windows or refrigerators, or taking off or putting on clothes which can affect body temperatures. Additionally, detection of drinking can differ for hot liquids, for which steam can rise in front of a person's face, affecting perceived temperatures. Such limitations will be addressed in future work.

4.3. Future Work

Next steps will focus on extending and improving PastVision+, as well as investigating other modalities for inference.

1. Improved segmentation of detected objects will enhance performance in challenging scenarios with warm background regions.
2. Body part heat trace dissipation will be explored over longer time periods and in various contexts, such as when people have been moving or idle.
3. Segmentation of multiple co-occurring heat traces and material-based modeling of sequential thermal touch activities will yield more powerful inference.
4. Tracking strategies will be used for detected objects, individuals, and heat traces in videos, also when a robot itself is moving.
5. Methods such as process models will be used for activity prediction based on thermovisual inference.
6. As well, additional evaluation of features, rules, and preprocessing will improve performance.
7. Problem cases and countermeasures will be identified, such as if a user touches more than one medicine bottle, more than one user is taking medicine, or one user is feeding another, like a parent and child; anomaly detection could be used to detect some unusual cases like putting a pill in one's mouth but then removing it; moreover, heat traces cannot be seen while a person holds an object; thus, ability to detect holding, possibly based on hand proximity and detecting if an object seems to be "floating," will be useful, in addition to detecting other potentially adversarial behavior, like hiding lips.
8. And, methods of inference targeting other routes of medicine administration will be explored.
9. As well, human science aspects such as the balance of camera resolution versus peace of mind can be investigated.

We believe there is much potential also in other modalities for useful inference.

1. Visually, using a UV camera lens will let robots assess skin health and evaluate if topical lotions have been applied correctly; radar will allow inference through walls and objects.
2. Aurally, much inference will be possible; e.g., laser microphones could be used to detect coughing remotely.
3. Not least, the amazing abilities of some animals to track smells and determine by smell "who has touched what" suggest that olfaction, although highly complex, will open the door to many possibilities. We believe that for example future systems will learn to identify individuals based on mixes of low-molecular-weight fatty acids contributing to body odor, and will also be able to track changes in compounds such as 2-nonenal to assess aging (Yamazaki et al., 2010).

In this way, by better being able to infer what humans have done, we believe robots will be better able to estimate how humans are feeling and what they are planning to do and will be able to use this knowledge to better care for humans, toward supporting their well-being.

AUTHOR CONTRIBUTIONS

MC came up with concept, implemented, tested, and wrote the article. JB contributed substantially to the concept and provided feedback for revising. Both authors have approved the article and agreed to be accountable for the work.

ACKNOWLEDGMENTS

We received valuable advice from our partners in the REMIND project including an experienced former medical doctor and would also like to thank the participants and all others who helped us.

FUNDING

The authors received funding from the Swedish Knowledge Foundation (Sidus AIR no. 20140220 and CAISR 2010/0271) and some travel funding from the REMIND project (H2020-MSCA-RISE No 734355).

REFERENCES

Abdelrahman, Y., Shirazi, A. S., Henze, N., and Schmidt, A. (2015). "Investigation of material properties for thermal imaging-based interaction," in *CHI 2015*, Seoul.

Benko, H., Wilson, A. D., and Balakrishnan, R. (2008). "Sphere: multi-touch interactions on a spherical display," in *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology (UIST 2008)*, Monterey, 77–86.

Clausius, R. (1854). Über eine veränderte form des zweiten hauptsatzes der mechanischen wärmetheorie. *Ann. Phys.* 93, 481–506. doi:10.1002/andp.18541691202

Cooney, M., and Bigun, J. (2017). "Pastvision: exploring "seeing" into the near past with thermal touch sensing and object detection–for robot monitoring of medicine intake by dementia patients," in *30th Annual Workshop of the Swedish Artificial Intelligence Society (SAIS 2017)*, Karlskrona.

Cooney, M., Nishio, S., and Ishiguro, H. (2015). Importance of touch for conveying affection in a multimodal interaction with a small humanoid robot. *Int. J. Hum. Robot.* 12, 1550002. doi:10.1142/S0219843615500024

Cover, T., and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Trans. Inform. Theory* 13, 21–27. doi:10.1109/TIT.1967.1053964

Crick, C., and Scassellati, B. (2008). "Inferring narrative and intention from playground games," in *Proceedings of the 12th IEEE Conference on Development and Learning*, Monterey.

Dragone, M., Saunders, J., and Dautenhahn, K. (2015). On the integration of adaptive and interactive robotic smart spaces. *Paladyn J. Behav. Robot.* 6, 165–179. doi:10.1515/pjbr-2015-0009

Graf, B., Reiser, U., Hägele, M., Mauz, K., and Klein, P. (2009). "Robotic home assistant Care-O-bot® 3-product vision and innovation platform," in *IEEE Workshop on Advanced Robotics and Its Social Impacts-ARSO 2009*, Tokyo, 139–144.

Gray, J., Breazeal, C., Berlin, M., Brooks, A., and Lieberman, J. (2005). "Action parsing and goal inference using self as simulator," in *RO-MAN 2005*, Nashville, 202–209.

Hotze, W. (2016). *Robotic First Aid: Using a Mobile Robot to Localise and Visualise Points of Interest for First Aid*. Master's Thesis, Halmstad University, Halmstad.

Kawamoto, H., Taal, S., Niniss, H., Hayashi, T., and Kamibayashi, K. (2010). "Voluntary motion support control of robot suit HAL triggered by bioelectrical signal for hemiplegia," in *Proceedings of IEEE Eng Med Biol Soc*, Buenos Aires, 462–466.

Kazemi, V., and Sullivan, J. (2014). "One millisecond face alignment with an ensemble of regression trees," in *CVPR 2014*, Columbus.

King, D. E. (2009). Dlib-ml: a machine learning toolkit. *J. Mach. Learn. Res.* 10, 1755–1758. doi:10.1145/1577069.1755843

Larson, E., Cohn, G., Gupta, S., Ren, X., Harrison, B., Fox, D., et al. (2011). "Heatwave: thermal imaging for surface user interaction," in *CHI 2011 Session: Touch 3: Sensing*, Vancouver.

Lundstrom, J., De Morais, W. O., and Cooney, M. (2015). "A holistic smart home demonstrator for anomaly detection and response," in *2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*. doi:10.1109/PERCOMW.2015.7134058

Mukai, T., Hirano, S., Nakashima, H., Kato, Y., Sakaida, Y. Guo, S., et al. (2010). "Development of a nursing-care assistant robot RIBA that can lift a human in its arms," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, Taipei, 5996–6001.

Noury, N. (2005). "AILISA: experimental platforms to evaluate remote care and assistive technologies in gerontology," in *Proceedings—HEALTHCOM 2005*, Busan, 67–72.

Osterberg, L., and Blaschke, T. (2005). Adherence to medication. *N. Engl. J. Med.* 353, 487–497. doi:10.1056/NEJMra050100

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.

Pollack, M., Engberg, S., Thrun, S., Brown, L., Colbry, D., Orosz, C., et al. (2002). "Pearl: a mobile robotic assistant for the elderly," in *Proceedings of the AAAI Workshop on Automation as Caregiver: The Role of Intelligent Technology in Elder Care*, Edmonton.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). "You only look once: unified, real-time object detection," in *CVPR*, Las Vegas.

Ruffell, A., and McKinley, J. (2005). Forensic geoscience: applications of geology, geomorphology and geophysics to criminal investigations. *Earth Sci. Rev.* 69, 235–247. doi:10.1016/j.earscirev.2004.08.002

Sagonas, C., Tzimiropoulos, G., Zafeiriou, S., and Maja, P. (2013). "300 faces in-the-wild challenge: the first facial landmark localization challenge," in *The IEEE International Conference on Computer Vision (ICCV) Workshops*, Sydney Vol. 117, 397–403.

Takacs, B., and Hanak, D. (2008). A prototype home robot with an ambient facial interface to improve drug compliance. *J. Telemed. Telecare*. 14, 393–395. doi:10.1258/jtt.2008.007016

Vernon, D., Thill, S., and Ziemke, T. (2016). "The role of intention in cognitive robotics," in *Toward Robotic Socially Believable Behaving Systems – Volume I Volume 105 of the Series Intelligent Systems Reference Library* (Cham: Springer), 15–27.

Vidas, S., Moghadam, P., and Bosse, M. (2013). "3D thermal mapping of building interiors using an RGB-D and thermal camera," in *IEEE International Conference on Robotics and Automation (ICRA)*, Karlsruhe, 2303–2310.

Yamazaki, S., Hoshino, K., and Kusuhara, M. (2010). Odor associated with aging. *Anti Aging Med.* 7, 60–65. doi:10.3793/jaam.7.60

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer, FB, and handling Editor declared their shared affiliation.

Copyright © 2017 Cooney and Bigun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.