Machine Vision and Applications manuscript No.
(will be inserted by the editor)

Body temperature estimation of a moving subject from thermographic images

Guillaume-Alexandre Bilodeau · **Atousa Torabi** · **Maxime Levesque** · **Charles Ouellet** · **J.M. Pierre Langlois** · **Pablo Lema** · **Lionel Carmant**

Received: date / Accepted: date

Abstract The continual measurement of the body temperature of a moving subject in a non-invasive way is a challenging task. However, doing so enables the observation of important phenomena with little inconvenience to the subject, and can be a powerful tool for understanding physiological reactions to diseases and medications. In this paper, we present a method to obtain the body temperature on a moving subject from thermographic images. The camera's output (a measurement for each pixel) is processed with a particle filter tracker, a clustering algorithm, and a Kalman filter to reduce tracking and measurement noise. The method was tested on videos from animal experiments and on a human patient. Tracking

G.A. Bilodeau · A. Torabi · C. Ouellet · J.M.P. Langlois
Department of Computer and Software Engineering,
École Polytechnique de Montréal,
P.O. Box 6079, Station Centre-ville Montréal (Québec), Canada, H3C 3A7,
E-mail: guillaume-alexandre.bilodeau@polymtl.ca

M. Levesque · P. Lema · L. Carmant
Centre de Recherche du Centre Hospitalier Universitaire Sainte-Justine, 3175, Côte Ste-Catherine, Montréal, (Québec), Canada, H3T 1C5

performance was then evaluated by comparison with manually selected regions of interest in thermographic images. The method achieves RMS temperature estimation errors of less than $0.1°C$.

Keywords Thermography · body temperature measurement · particle filter tracker · Kalman filter

1 Introduction

Continual measurement of the body temperature of a moving subject in a non-invasive way is a challenging task. However, doing so makes it possible to observe important physiological phenomena with little or no inconvenience to the subject, and can be a powerful tool for understanding reactions to diseases, medication, and stress.

Thermography has had significant medical applications for more than a decade [12]. The range of applications is very broad and includes aspects of avian flu [4, 7], dermatology [19, 15], detection of eye abnormalities [1, 17], breast cancer [3, 18] and stress response [21]. Sudden infant death syndrome may be detected based on an increase in facial temperature [10, 20]. In most applications, the focus is on analyzing still images, or in analyzing images separated by long intervals. For example, Sunderam et al. [22] obtained thermal images of the faces of six patients every hour and during seizure events as indicated by real-time EEG analysis. The images were manually filtered to remove any where occlusions had occurred. Since the face was in the middle of the image in these cases, the measured temperature was maximal in the center region and there was no tracking requirement. In other cases, either the subject is not moving significantly with respect to the camera [21], or special markers are placed on the body to aid in tracking [16].

The present work was developed in the context of studies on epilepsy [14,6]. Sunderam et al. [22] have shown that seizures may affect body temperature, and Ahlenius et al. [2] have shown that kainic acid (KA), a drug used in animal models of temporal lobe epilepsy, impacts the body temperature of a laboratory rat. Thus, body temperature could be used to detect the occurrence of seizures, and, in general, monitoring the body temperature has the potential to enhance research on epilepsy and seizures. Our interest therefore lies in continually monitoring the temperature of an area on the body of a subject that can move inside a perimeter within the field of view of a thermographic camera. A thermographic camera gives as output an image where each pixel corresponds to a temperature measurement in its field of view. To measure the temperature of the same area in a video sequence, that area in the image and the average temperatures (pixel values) inside it need to be tracked.

Our proposed method is based on a particle filter that tracks a user-specified region in a stream of images. The tracked region is then analyzed to remove outlier measurements in the region (e.g. fur or hair) using simple clustering. Finally, outlier measurement regions are filtered using a Kalman filter which estimates the body temperature by filtering measurement noise. We considered two applications. The first involves laboratory rats presenting seizures due to brain lesions, neonatal hyperthermia, or the injection of KA. In this case, temperature measurements are used to confirm seizure occurrences detected with local field potentials (LFP). The second involves the measurement of the temperature of the face of a human patient. An early version of this work has already been presented [23].

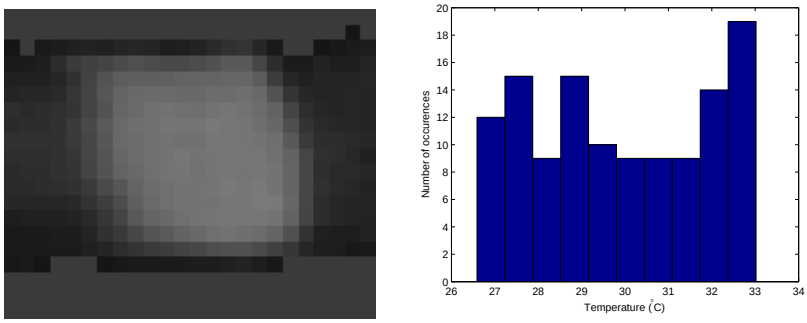
The paper is structured as follows. Section 2 gives a precise definition of the problem. Section 3 presents our measurement methodology. The acquisition setup for humans and animals and our experimental results are presented and discussed in section 4. Section 5 concludes the paper.

2 Problem definition

The problem considered in this paper is the continuous monitoring of the surface temperature fluctuations of the body of a moving subject in the least invasive way possible. Using a thermographic camera for this purpose presents distinct opportunities. The camera can be used without inconveniencing the subject, and it is clearly well suited to continuous monitoring. Moreover, it is ideal for use in the case of laboratory animals and other non-collaborating subjects, since it eliminates the need for an invasive sensor, which the subject may find uncomfortable and attempt to remove, and its tether, which may become tangled during the subject's movements.

However, using a thermographic camera for this task also poses significant challenges:

1. Each pixel of an image obtained with a thermographic camera corresponds to a temperature measurement. The measured temperatures for each pixel are typically different for a body area that may include hair and fur. This is shown by the intensity (temperature) histogram of a shaved body area of a rat in Fig. 1. The lighter pixels correspond to exposed skin (note that the measured temperature is not uniform), while the darker pixels correspond to fur. The mid-gray pixels correspond to background. Thus, when measuring a body area, a strategy needs to be devised to combine many measures into a single one for the whole body area. Also, measurements corresponding to fur, hair, and background should be filtered out.
2. In the present project, we assume that the subject is always located in the camera's field of view, but it is free to roam inside a given perimeter. Since different body parts have different surface temperatures, e.g. the temperature of a rat's ears, fur, skin, and tail are all different and are related in different ways to the animal's core temperature, we need to always measure the temperature in the same area. Furthermore, a target body part



(a) (b)

Fig. 1 Portion of a skin region. a) intensities, b) temperature histogram

may become occluded owing to the subject's movements or body position, or to the presence of objects in its environment. Thus, the area to be measured must be visible most of the time and it must be tracked. Skin region are a better choice for observing temperature variations, because fur and hair are thermal insulators. Fig. 2 illustrates different measurement cases. Fig. 2 a) and Fig. 2 b) show the mean temperature of the rat's head and of a shaved patch on its body. Because the head is occluded by an LFP measuring device, its visibility changes resulting in large temperature variations (32.76 to 32.34 $^{\circ}C$). Fig. 2 c) shows that the surface area of hot regions is not a good criteria for temperature measurements, because of possible occlusions that may reduce their size or even split them in two, as illustrated. Furthermore, Fig. 2 d) shows that other hot distractor objects (for example, the hands or body parts of experimenters) can move into the field of view of the camera. Therefore, a formal tracking method is required to measure always temperature on the same body area.

The focus of the present study was to continuously and automatically monitor core temperature fluctuations, as estimated from surface temperature measurements of the subject. The surface temperature is assumed to be highly correlated to the subject's core temperature.

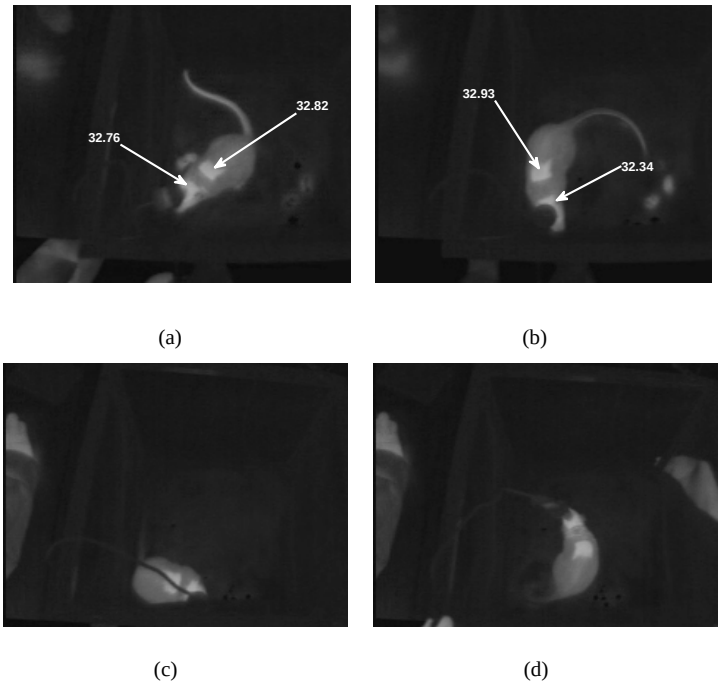


Fig. 2 Different measurement cases. a) and b) body and head mean temperature ($^{\circ}C$), given head visibility, c) change in size of measuring area (split in two), and d) distractors (hands around the cage).

In the context of neurophysiological studies, the absolute value of the surface temperature, as measured by the thermographic camera, is therefore inconsequential. What matters is its evolution in time as the subject carries on with its normal activities or as its pathological condition persists or changes. Every pixel of an image obtained with a thermographic camera corresponds to a temperature measurement. The goal of this project is thus to select temperature measurements (pixels) that will allow temperature variations to be monitored. We must select pixels that are always on the same body area and that correspond to skin. This is the focus of the methodology presented in this work.

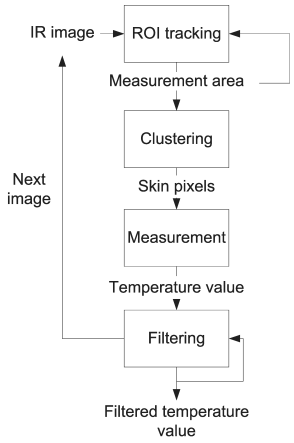


Fig. 3 Overview of the measurement method

3 Methodology

The proposed method processes a video sequence and outputs the mean of the temperature measurements for each frame in a tracked area. Fig. 3 gives an overview of the method. During initialization, control points that lie on the boundary of a region of interest (ROI) are selected in a frame. The ROI is located in subsequent video frames using a particle filter [11]. The ROI should contain exposed skin, but it can also include hair or fur. Skin is separated from other types of surfaces using temperature clustering. Pixels corresponding to skin are placed in the warmer cluster (higher pixel intensity). In some cases, the ROI is not located correctly, which results in an incorrect temperature measurement. To filter incorrect mean temperatures, we use a Kalman filter [13] to estimate the temperature curve and remove outliers. This whole process results in a curve displaying the evolution of the mean temperature of the ROI over time.

3.1 Particle filter tracking

We impose the requirement that the tracking algorithm must not lose the location of the ROI for the duration of the video, and it must not be distracted by other hot regions which are not at exactly the same temperature (e.g. head, eyes). We have observed that the temperature distribution across the surface of the body is rarely uniform. We applied a particle filter tracker [11] adapted to this tracking problem. Particle filter tracking is robust to sudden movements and occlusions in long duration video sequences, and so is appropriate for our conditions.

Our tracking algorithm is based on the following assumptions:

- the images are grayscale, with white (intensity of 255) meaning hot and black (intensity of 0) meaning cold in a given range;
- the ROI is approximately elliptical in shape, corresponding to a shaved body area on an animal or on a human face;
- the temperature of the ROI on the subject is higher than its surroundings;
- the ROI can be temporarily occluded by medical equipment, e.g. by a Local Field Potential (LFP) recording device and its associated wiring; and
- the ROI is always in the camera's field of view.

The particle filter is a Bayesian tracking method, which, in each time t, recursively approximates the state of the tracking target as a posterior distribution using a finite set of weighted samples. A sample is a prediction about the state of the tracking target. In our case, it is a candidate area in the image.

In this work, we propose to model the tracking target by an ellipse fitted to control points which lie on the edges of an ROI's perimeter (Fig. 4). We selected an ellipse as a model because it is simple, but flexible enough to approximately fit many types of boundaries.

(a) (b)

Fig. 4 Ellipse fitted to control points on: a) a shaved body area perimeter, and b) around a face.

Good fitting of boundaries allows more precise localization of the ROI. The ellipse is defined by

$$Ax^2 + Bxy + Cy^2 + D = 0; \tag{1}$$

The parameters A, B, C, and D are computed to optimize ellipse fitting. The particle filter state $\mathscr{X}_t$ at each time step t is defined as a set of N vectors $\mathbf{X}_t^j$ of control points, where each control point is represented by a pair of x and y coordinates, that is $\mathscr{X}_t = \{\mathbf{X}_t^j | j = 1, \ldots, N\}$. We used intensity and edge features as the observation model for our particle filter tracking method. The observation model is used to determine the similarity likelihood (weight) of the samples with the target area. The edge feature is used to align the ellipse on the boundary of the ROI, and the intensity feature is used to align the ellipse on an ROI with the same intensity (or temperature) as the model.

For the edge feature, we use the dot product between a unit normal vector of the ellipse at the positions of control points and the image gradient at the same positions as a measure. The dot product is used, because a large gradient magnitude around the ROI's perimeter is desirable and the gradient direction should be perpendicular to the perimeter. The gradient measure $\phi_g(s)$ for a given sample s (a candidate ROI in the image) is

$$\phi_g(s) = \frac{1}{N_{cp}} \sum_{k=1}^{N_{cp}} \left| \mathbf{n}^k(s) \cdot \mathbf{g}^k(s) \right|, \tag{2}$$

where $\mathbf{n}^k(s)$ is the unit normal vector of the ellipse at the k^{th} control point, $\mathbf{g}^k(s)$ is the orientation of the image gradient at the same point, and N_{cp} is the number of control points. To facilitate the fusion of edge and intensity features, the gradient measure is normalized by subtracting the minimum gradient measure of the sample set S_t (the set of all candidate areas in an image) and dividing by the gradient measure value range, as follows:

$$\bar{\phi_g}(s) = \frac{\phi_g(s) - min_{s^j \in S_t}(\phi_g(s^j))}{max_{s^j \in S_t}(\phi_g(s^j)) - min_{s^j \in S_t}(\phi_g(s^j))}. \tag{3}$$

For the intensity feature, the Euclidean distance between the average intensity μ_s of the sample ellipse pixels and the average intensity μ_r of the template ellipse pixels is used as a measure. The template ellipse is the result of ellipse fitting to the chosen control points in the first frame of the video sequence. The intensity measure $\phi_i(s)$ for a given sample s is

$$\phi_i(s) = |\mu_s - \mu_r|. \tag{4}$$

The intensity measure is normalized as follows:

$$\bar{\phi_i}(s) = \frac{max_{s^j \in S_t}(\phi_i(s^j)) - \phi_i(s)}{max_{s^j \in S_t}(\phi_i(s^j))} \tag{5}$$

At time t, samples are selected with replacement from the sample set

$$S_{t-1} = \left\{ \mathbf{X}_{t-1}^j, w_{t-1}^j \right\}_{j=1}^{N}, \tag{6}$$

where N is the number of samples, $\mathbf{X}_{t-1}^j$ is the vector of control points of the j^{th} sample at time $t-1$, and w_{t-1}^j is its corresponding weight. Sample set S_{t-1} is the approximation of the

target state posterior distribution at time $t-1$. N_{s_j} samples are chosen with a probability of w_{t-1}^j (the weight), which is derived for sample j by

$$w_{t-1}^j = \bar{\phi}_g(s_{t-1}^j) + \bar{\phi}_i(s_{t-1}^j). \tag{7}$$

This means that samples with large weights may be chosen several times (i.e. identical copies of one sample), and samples with small weights may not be chosen at all. In two consecutive frames, the particle filter state does not change dramatically. It is mostly a translation of the ellipse's major and minor axes and rotation around its center. In each time step t, samples are propagated in the state space using a first-order auto-regressive dynamic model defined as

$$\mathscr{X}_t = \mathscr{X}_{t-1} + \omega_t, \tag{8}$$

where $\mathscr{X}_t$ and $\mathscr{X}_{t-1}$ are particle filter states at time t and $t-1$ respectively, and ω_t is the stochastic part of the dynamic model. It is a vector of multivariate Gaussian random variables and corresponds to random translations and rotations of the ellipses. For the temperature measurement, we use the best sample s_b, which is

$$s_b = argmax_{s^j \in S_t} \left\{ \bar{\phi}_g(s^j) + \bar{\phi}_i(s^j) \right\}. \tag{9}$$

The algorithm is as follows:

1. Initialization. Manually select control points on the perimeter of the ROI.

For each new frame:

2. Threshold the infrared image to keep the hottest regions (pixels with an intensity greater than 80). Find the gradient using a Sobel edge detector [8].
3. Select samples from the sample set S_{t-1} based on the weights (Equation 7).

4. Propagate samples using Equation (8).
5. Fit an ellipse to the control points of each sample.
6. Compute the observation measures to update new sample weights using Equation (7).
7. Choose the best sample using Equation (9).

3.2 Extracting skin pixels and estimating the temperature

The area for measuring the temperature is bounded by the ellipse fitted to the control points of s_b. The particle filter's best sample is a region that may include surfaces other than skin (e.g. fur, hair, clothes, medical devices, or cabling, etc.). Furthermore, the temperature across the skin region is not uniform (see Fig. 1). The temperature is estimated by calculating the average of a group of pixels in order to increase robustness in the presence of noise. To arrive at an accurate temperature measurement, pixels other than skin must therefore be removed before calculating the average. This pixel removal and the selection of the most representative pixels is not a trivial problem, since the measurement area is not uniform.

To obtain stable results, we have found that calculating an average over P% of the pixels in the area inside the best sample, followed by a filtering method, leads to stable and repeatable results. Thus, to find the best skin pixels for measurement, we clustered the data in two groups, the first composed of the hottest (highest intensity) P% of the pixels, and the second containing the remaining pixels. Only the first group of pixels is used to evaluate the temperature. The intensity C_{s_b} is the mean of the intensities of these pixels.

Thus, the temperature $T_{s_b}(f)$ for the best sample in frame f is calculated by

$$T_{s_b}(f) = T_{min} + \left((C_{s_b}/255) * (T_{max} - T_{min})\right), \tag{10}$$

where T_{min} and T_{max} are the minimum and maximum values of the temperature range selected for the camera, and C_{s_b} is the mean intensity of the pixels selected by clustering in

the area of the sample s_b. Note that the camera was set to have a linear relationship between intensity and temperature, and relating intensity to temperature is the reason why Equation 10 is linear.

We have chosen to estimate temperature using the mean of the P% hottest pixels in the area inside the best sample for two reasons: 1) since measurements may be noisy, taking only a few pixels will result in noisy results, and 2) averaging over all the pixels will include pixels that do not belong to skin. The value of P is a compromise between filtering noise and not including non-skin pixels. Typically, localization errors by the particle filter are less than 2 pixels in x and y 66% of the time, and less than 5 pixels 99% of the time for a body area of 10×10. This is the case in two of the rat experiments. For the other experiments, the body area is 40×40 with localization errors between 5 and 10 pixels. For the 10×10 case, a shift of the center of the best sample by 5 pixels in x and y gives a surface intersection of approximately 25%. Thus, fixing P at 25% allows us to average the measurements of only skin pixels 99% of the time. This rationale also applies for the larger body area.

3.3 Temperature filtering

To compensate for camera noise that may perturb temperature measurements, the temperature values obtained at every frame are filtered using a $1D$ Kalman filter [9]. The Kalman filter prediction is

$$\bar{x}_i^- = d_i \bar{x}_{i-1}^+ \tag{11}$$

$$\sigma_i^- = \sqrt{\sigma_{di}^2 + d_i(\sigma_{i-1}^+)^2} \tag{12}$$

In our process model, we assume that the temperature is constant over two consecutive measurements. The dynamic model state change matrix d_i thus reduces to the scalar 1 in Equation 11. Equation 12 calculates the expected standard deviation of the whole process (system noise + measurement noise) based on the measurement noise. We assume that the subject's body (system) generates a temperature with a standard deviation of 0.2 (σ_{di}=0.2 in Equation 12, a relatively large noise), since we do not know the true body "noise" (this is explained in section 4.3).

The Kalman filter correction is

$$x_i^+ = \left(\frac{\overline{x_i^-} \times (\sigma_{mi})^2 + m_i T_{s_b}(f) \times (\sigma_i^-)^2}{(\sigma_{mi})^2 + m_i^2 (\sigma_i^-)^2} \right) \tag{13}$$

$$\sigma_i^+ = \sqrt{\left(\frac{\sigma_{mi}^2 (\sigma_i^-)^2}{\sigma_{mi}^2 + m_i^2 (\sigma_i^-)^2} \right)}. \tag{14}$$

For the correction, we assume that the temperature is observed (m_i=1) and its standard deviation is calculated from V temperature values ($\sigma_{mi} = std(V)$). The final temperature value for a frame is x_i^+ (Equation 13). The process standard deviation is updated in Equation 14.

For our Kalman filter, we need to estimate the camera measurement noise σ_{mi} using V temperature measurements. This is explained in section 4.1.

4 Experimentation

We now present the experimental methodology, results, and discussion. The measurement methodology was tested on rats and on a human face.

4.1 Data acquisition and experimental setup

The temperature sensor is a Thermovision A40M thermographic camera (FLIR Systems). From the manufacturer's specifications, the accuracy is $\pm$ 2%, but the precision (or measurement noise) is not specified. The camera was configured with a linear measurement range of $20°C$ to $40°C$, and the pixels were quantified with 8 bits. That is, pixel values of 0 and 255 corresponded to temperatures of $20°C$ or lower and $40°C$ or higher respectively. The temperature interval between two adjacent pixel values is therefore $0.078°C$. Without averaging a pixel region, the precision should be half this interval, that is $\pm$ $0.039°C$. By averaging over a region, we may obtain slightly better precision.

To evaluate the precision of the camera, data that do not include tracking errors are required to prevent over-smoothing and the loss of small temperature variations. To obtain the V temperature measurements needed, the temperature of a formica tabletop was estimated by calculating the mean of the 20 hottest pixels in the ROI. We selected this object because its emissivity (0.94) is similar to that of skin (0.98). We captured thermographic images of the tabletop from a fixed point of view at 27 frames/s for approximately 30 minutes. The tabletop was located in a room where the temperature was controlled by an electronic thermostat set to approximately $24°C$. We selected an ROI of 20×20 pixels in the middle of the image. Assuming that the temperature changed smoothly, and because the room temperature was not controlled, we computed the regression of the data, using a 7^{th} order polynomial, and the average fitting error. Fig. 5 shows the mean measured temperature and the fitted polynomial. The camera measurement noise, which is, the average fitting error, was calculated at $0.021°C$, with a standard deviation of $0.026°C$. σ_{mi} was thus set to $0.021°C$.

We did not validate the accuracy of the camera, since we did not have the equipment to do so. However, in a previous work with the same camera, the accuracy was evaluated at

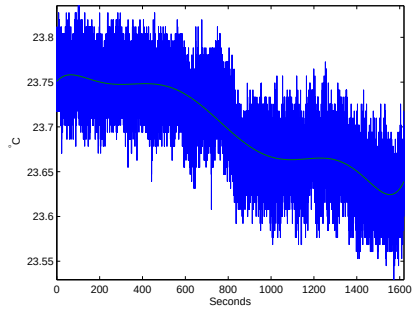


Fig. 5 Mean temperature measured for a tabletop during a 1620 second period, and the fitted polynomial to evaluate the camera measurement noise.

Fig. 6 Experimental setup. Camera and mesh cage

$0.13°C$ [4] between $35°C$ and $38°C$. In the same study, the drift was said to be negligible during experiments lasting as long as 8 hours.

4.1.1 Data acquisition for the rat videos

During the animal experiments, a Sprague-Dawley rat (Charles River Laboratories) was placed in a cubic metal mesh cage with an open top (see Fig. 6), and the thermographic camera was angled down toward the cage. The camera was mounted on a 525MV tripod (Manfrotto) and pointed toward the open top of the cage at an angle of approximately 20° with the vertical. The animal was 1 meter away from the camera, which was configured with a linear measurement range and the appropriate emissivity of the subject.

During the initial experiments, we determined that the rat's fur prevented precise measurements of its body temperature. Our solution was to shave an area of approximately 10 cm^2 on the rat's body to provide direct visibility to the animal's skin. The rat's head, an interesting region that is less covered in fur, was not always visible and was often occluded by an LFP recording device (Neuralynx Cheetah System) connected to electrodes inside the rat's brain. Observing the rat from the top, a shaved area on its back provided better results, since this region is almost continually visible because the rat tends to remain on its four feet. The shaved area was approximately elliptical in shape. Fig. 4(a) shows a typical frame that must be processed.

Our animal experiments consisted of five videos. Their technical descriptions are presented in Table 1. The videos titled Rat1 and Rat2 feature different rats injected with 6 mg/kg i.p. of kainic acid (KA) (Sigma-Aldrich Canada Ltd) during two-hour experiments. A wide angle lens was used, resulting in a smaller measurement area than for videos Rat3 to Rat5. The videos titled Rat3 and Rat4 feature a third rat with spontaneous seizures in two experiments, each lasting approximately 25 minutes. The standard camera lens was used. The video titled Rat5 features a fourth rat, filmed for approximately 16 minutes shortly after anesthesia. The standard camera lens was used. All experimental procedures conformed to institutional policies and guidelines (Ste. Justine Research Center, Université de Montréal, Québec, Canada). All videos had a 320×240 resolution and were compressed with an Xvid FFDshow encoder (Quality: 100%) (http://sourceforge.net/projects/ffdshow). The acquisition frame rate depends on the computer configuration and hard drive speed. If the computer is not fast enough, frames are dropped.

Table 1 Description of experiments

Video name	Duration (s)	Number of frames	Frame rate (frame/s)	Approximate size of body area to measure
Rat1	7037	90705	12.89	10×10
Rat2	7094	208668	29.41	10×10
Rat3	1394	40999	29.41	40×40
Rat4	1045	30745	29.41	40×40
Rat5	936	27162	29.01	40×40
Human1	1126	33132	29.41	40×60

4.1.2 Data acquisition for human face video

Our method was designed primarily to process rat videos, but can be generalized to other kinds of subjects. We acquired thermographic images of a human face by filming a subject working on a computer (see Fig. 4(b)). The camera was located on a desk in front of the subject. The experiment lasted approximately 19 minutes (see Table 1 for details). The standard camera lens was used. The camera was configured with a linear measurement range, with the appropriate emissivity for the subject.

4.2 Evaluation methodology and metrics

Our method was implemented in Matlab (The MathWorks) and run on a Xeon 5150 2.66 *GHz* computer (Intel). For particle filter tracking, eight control points were used for ellipse fitting. The points were chosen to cover the entire ROI perimeter, and the distance between them was approximately the same. In our experiments, the ROI translational motion had a higher frequency than the rotational motion. Consequently, 75% of the samples were specified for random translational motion and 25% of the samples for random trans-

lational+rotational motion. Our tracking algorithm was tested by varying the number of samples N (50, 100, or 200 samples), that number being chosen arbitrarily to show its effect on the tracking results and processing time. Temperature calculations were based on the $P = 25\%$ hottest pixels in the tracked area, as described in the methodology (see section 3.2).

Since we collected a large amount of data, we used two metrics to measure the performance of our tracking algorithm. We generated a basis for comparison by randomly selecting F frames ($F = 650$) over the whole video sequence. It is important to note here that we obtain temperature values from the pixels of a thermographic image within an ROI. Our basis for comparison therefore uses these pixels, and not the actual temperature of the subject, as we are not trying to evaluate the accuracy of the camera. We manually selected the center of the shaved area and of the forehead for the rat and human videos respectively. A bounding polygon was built, and the temperature values were calculated using equation 10. The center of the shaved area and the forehead were also recorded for localization error assessment. This gave us a set of manually selected ROI temperatures T_{ROI} and a set of manually selected localizations L_{ROI}. The localization errors of the tracker were evaluated using the mean absolute distances in x and y and their standard deviation. For x, they are defined by

$$\mu_x = \frac{1}{F}\sum_{i=1}^{F}\left|L_{s_{b_x}}(i) - L_{ROI_x}(i)\right|, \tag{15}$$

and

$$\sigma_x = \sqrt{\frac{1}{F}\sum_{i=1}^{F}\left|L_{s_{b_x}}(i) - \mu_x\right|}, \tag{16}$$

where $L_{s_{b_x}}(i)$ is the location in x of the center of the best sample s_b. The equations are similar for y.

The temperature measurements $T_{s_b}(i)$ obtained by the tracking algorithm for these frames were also compared with the manually selected ROI, because a good temperature measurement can be obtained even if the location of s_b is not perfect. The evaluation metric is the root mean square (RMS) error, defined as

$$T_{rms} = \sqrt{\frac{1}{F}\sum_{i=1}^{F}(T_{s_b}(i) - T_{ROI}(i))^2}. \quad (17)$$

4.3 Results and discussion

4.3.1 Tracking precision for rat videos

Table 2 gives the localization errors and RMS errors obtained for the rat videos using different numbers of samples N, with and without the Kalman filter. Results show that a larger number of samples reduces the localization errors in general. We say "in general" because the stochastic nature of particle sampling gives different results for each run (see Equation 8). The mean localization errors of many runs are smaller with more samples, which is because, with more samples, the tracking area is located more precisely. A larger number of samples is particularly useful for videos in which the tracked target makes many sudden and large movements. For these cases, a larger coverage of the image area is necessary to estimate the most likely target position. The Rat1 and Rat2 videos are more difficult to process because of the large number of seizures experienced by the animal. Although the mean errors in x and y are around 1-2 pixels, they are proportionally larger relative to the size of the measurement area, which is small. Thus, occasional errors of 4-5 pixels give erroneous measurements because they are outside the measurement area. In this case, there is a large difference in performance between 50 samples and 100 samples. With 100 samples, the mean localization error and standard deviation are smaller, and thus the results for tem-

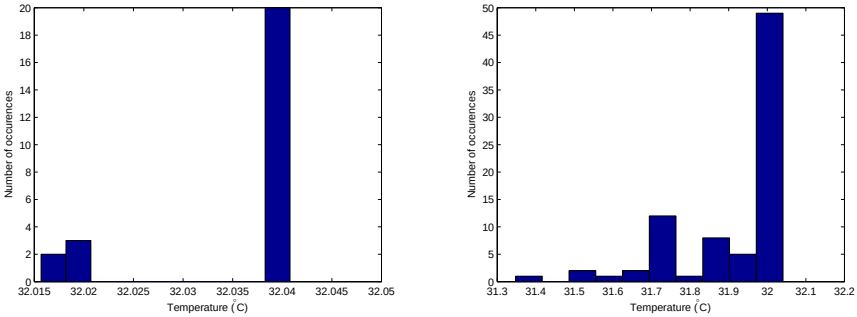
perature measurement will be more stable and should be considered a better choice than the use of 50 samples. With a small number of samples, samples are not distributed sufficiently evenly in all directions around the tracked target. Consequently, there is a risk of tracker failure owing to sudden movements of the rat's body. Using 200 samples provided no noticeable improvements in the results, but increased processing time.

The localization errors in x and y present only one aspect of the tracking performance. Since temperature measurements are obtained from the mean of a portion of the pixels inside a measurement area, a localization error of 5 pixels can give the same results for the resulting estimated temperature. Fig. 7 illustrates this with a plot of the number of times a mean temperature is obtained with Equation 10 for shifts of a measurement area smaller than or equal to ± 2 pixels and smaller than or equal to ± 4 pixels. As shown in the figure, an error of ± 2 pixels for localization gives approximately the same temperature (a $0.03°C$ difference). For a larger localization error of ± 4 pixels, there is a large variation in the temperature obtained (around $0.7°C$). This explains why, even though we obtain localization errors of between 5 and 12 pixels for the Rat3, Rat4, and Rat5 videos, we actually have a smaller RMS error on the temperature than for the Rat1 and Rat2 videos. The localization error is large in absolute terms, but small relative to the size of the measurement area (shaved patch). Thus, localization is important, but, because we use a portion of the pixels to estimate temperature, a small localization error has a limited impact on the resulting temperature.

By comparing the RMS error values for the five videos and 100 samples, we note that it varies between $0.06°C$ and $0.11°C$ for a body temperature of approximately $30°C$. This means that the noise level caused by tracking is approximately $0.1°C$. The pixel selection method is the same for our experiments and for the manually selected ROI, thus it is not part of the $0.1°C$ tracking error.

Table 2 Localization errors and root mean square error for each rat video. *N*: Number of samples, μ_x: mean localization error in x, σ_x: standard deviation of the localization error in x, μ_y: mean localization error in y, σ_y: standard deviation of the localization error in y, T_{rms}: root mean square error of temperature, FPS: frames processed per second.

Video	N	$\mu_x(\sigma_x)$	$\mu_y(\sigma_y)$	$T_{rms}(^\circ C)$	Kalman filter	Time (s)	FPS
Rat1	50	1.86 (2.37)	2.02 (2.75)	0.131	not active	19989	4.53
	50			0.120	active	20304	4.47
	100	1.77 (1.93)	1.63 (1.21)	0.085	not active	32300	2.81
	100			0.079	active	33113	2.74
	200	1.66 (2.03)	1.77 (1.42)	0.089	not active	67909	1.34
	200			0.083	active	69004	1.31
Rat2	50	0.87 (0.75)	1.53 (0.75)	0.157	not active	46979	4.44
	50			0.144	active	48096	4.34
	100	0.84 (0.61)	0.61 (0.72)	0.111	not active	85538	2.44
	100			0.106	active	87348	2.39
	200	0.81 (1.11)	0.56 (0.77)	0.110	not active	130654	1.60
	200			0.104	active	132875	1.57
Rat3	50	11.51 (7.92)	2.44 (1.86)	0.069	not active	9666	4.24
	50			0.064	active	9854	4.16
	100	8.99 (7.05)	7.05 (1.19)	0.072	not active	17387	2.36
	100			0.070	active	17796	2.30
	200	10.74 (5.16)	3.69 (1.61)	0.077	not active	31256	1.31
	200			0.076	active	31912	1.28
Rat4	50	5.02 (5.73)	4.18 (5.73)	0.076	not active	7120	4.32
	50			0.071	active	7294	4.21
	100	4.71 (5.23)	4.05 (3.72)	0.063	not active	12547	2.45
	100			0.062	active	12831	2.40
	200	4.63 (4.94)	4.26 (3.56)	0.064	not active	23666	1.30
	200			0.062	active	24109	1.28
Rat5	50	4.05 (4.18)	4.71 (5.62)	0.034	not active	5892	4.61
	50			0.032	active	5960	4.56
	100	3.22 (3.33)	3.24 (3.86)	0.056	not active	9865	2.75
	100			0.055	active	10005	2.71
	200	4.11 (3.41)	4.42 (3.93)	0.031	not active	16589	1.64
	200			0.030	active	16934	1.60



(a) (b)

Fig. 7 Number of occurrences of a mean temperature for different localization errors for frame 10420 of the Rat1 video. a) smaller than or equal to ± 2 pixels, b) smaller than or equal to ± 4 pixels

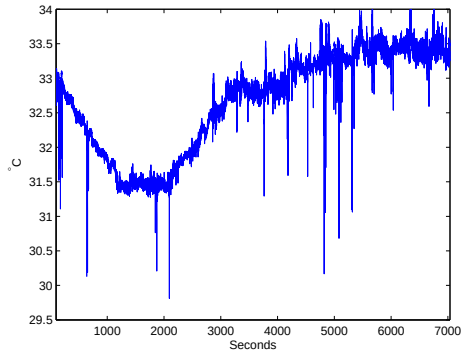
Fig. 8 illustrates a possible cause of tracking errors. The global decrease in temperature between 0 and about 1800 seconds is caused by the rat's physiological reaction to kainic acid (KA). This phenomenon was previously observed with a rectal thermometer at 15-30 minutes intervals [2]. By comparing Fig. 8(a) and Fig. 8(b), we note that the measurements with the tracker are noisier than for the manually selected ROI. Sudden temperature drops of more than $1^{\circ}C$ are caused by tracking errors. In this experiment, occlusion by the LFP recording device caused the tracking errors at around 150, 1850, 2100, 4850, and 5100 seconds. In such cases, only a portion of the shaved patch is visible. During these times, particle filter tracking cannot track the patch correctly, because of considerable changes in the patch shape and distraction by the rat's head. The other reason for the tracking errors is that the rat is moving too fast (e.g. wet dog shakes during a seizure) and because of large displacement in the images at around 3800, 4200, 4550, 5350, 6000, 6050, and 6700 seconds. Fig. 8(c) shows that, after 3000 seconds, the tracking errors are mostly positive because of the rat's quick movements. This resulted in tracking of the rat's head, which is slightly warmer than the shaved patch. It is noticeable that some negative errors are not represented on this graph, since the basis for comparison uses 650 randomly selected frames

and does not include all the tracking errors. However, it does indicate that negative errors are less frequent than the positive errors caused by incorrectly tracking the head. Negative errors mean that tracking is neither on the head nor on the patch.

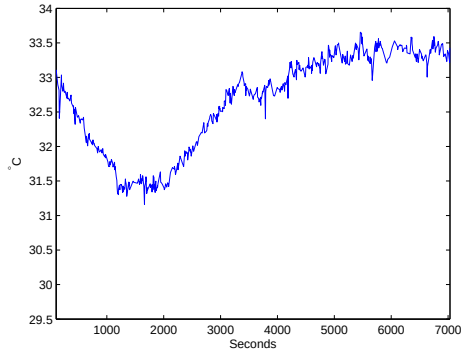
4.3.2 Benefits of the Kalman filter

As for the benefits of the Kalman filter, the results in Table 2 show that it reduces the RMS error. That is, it smooths temperature variations by removing estimated camera measurement noise. To be efficient, the Kalman filter needs a good estimate of the measurement noise. To remove measurement errors related to sensor noise, the standard deviation of the measurement noise σ_{mi} was estimated from the tabletop experiment, because the selected formica tabletop has an emissivity similar to that of skin (0.94 vs. 0.98). This allowed us to accurately estimate the measurement noise of the thermographic camera and remove this noise from our temperature results.

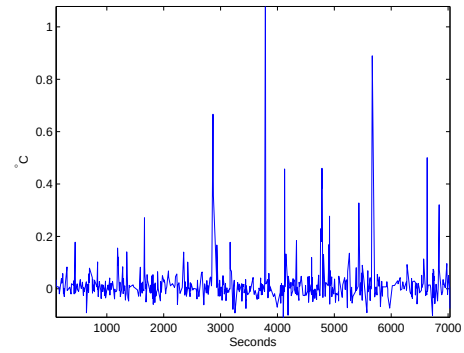
The Kalman filter could also be used to smooth the tracking localization errors by using instead data from actual experiments. To do this, we estimated σ_{mi} based on V temperature measurements. Table 3 shows the RMS error for the Rat4 video and different values of V. We also included the σ_{mi} calculated from the tabletop. With a large V, the RMS error is smaller. More measurements are used, and this means that we will not only estimate measurement noise, but also tracking noise, since tracking errors will eventually occur (σ_{mi} increases). However, this filter might smooth the data too much and remove small, local, and interesting temperature changes that are larger than the measurement noise, but smaller than the tracking and measurement noise combined. So, although the results might be better globally, the Kalman filter might remove interesting local information. So, we believe it is better to use a smaller V to estimate only the measurement noise. This is why, in our experi-



(a)



(b)



(c)

Fig. 8 Results from our tracker for Rat1 video compared to manually selected ROI ($N = 100$ and $V(\%) = 1$). a) Temperature values obtained with our particle filter tracker. b) Manually selected ROI temperature values in $F = 650$ frames. c) Error between the two.

Table 3 Root mean square error for the Rat4 video with different values of V ($N = 100$). V: number of temperature values for measurement noise, T_{rms}: root mean square error.

$V(\%)$	V	σ_{mi}	$T_{rms}(^{\circ}C)$
(Tabletop)	-	0.021	0.062
0.5	153	0.026	0.062
1	307	0.029	0.062
5	1537	0.040	0.061
10	3074	0.042	0.061
20	6149	0.107	0.058
50	15372	0.146	0.057
100	30745	0.122	0.058

ments, we use the σ_{mi} calculated from the tabletop. In any case, the choice of V depends on the window (in seconds) of the temperature changes in which the user is interested.

4.3.3 Tracking precision for the human video

Table 4 shows the results of the Human1 video. First, we obtain a similar RMS error to that for the rat, and with similar remarks. The worst result is with $N = 200$ samples, because of the randomness of the particle filter. On average, with many runs, we should obtain a better result with 200 samples. For localization errors, we obtain errors of about 9 pixels in x and y, but, since the ROI is large, the RMS error stays low.

4.3.4 Discussion

The RMS error obtained for all our experiments is larger than the camera measurement noise obtained from the tabletop (see section 3.3). There are two reasons for this: 1) tracking errors, and 2) the observation angle of the surface. Indeed, for the latter, we hypothesize that the target measurement area is deformable, and its normal is not always aligned with the

Table 4 Root mean square error and fitting error for the Human1 video. N: Number of samples, T_{rms}: root mean square error, μ_x: mean localization error in x, σ_x: standard deviation of the localization error in x, μ_y: mean localization error in y, σ_y: standard deviation of the localization error in y, FPS: frames processed per second.

Video	N	$T_{rms}(^{\circ}C)$	$\mu_x(\sigma_x)$	$\mu_y(\sigma_y)$	Kalman filter	Time (s)	FPS
Human1	50	0.088	8.82 (5.21)	8.63 (5.69)	not active	7236	4.58
	50	0.081			active	7298	4.54
	100	0.086	8.08 (5.21)	9.33 (5.72)	not active	12193	2.72
	100	0.085			active	12564	2.64
	200	0.103	9.44 (4.62)	9.13 (5.75)	not active	21763	1.52
	200	0.097			active	22302	1.49

camera sensor's normal. Thus, the infrared radiation measured by the camera may change with the angle of the target measurement area [5], which is why we use σ_{di}=0.2 (a large noise assumption) to try to filter this possible system noise. This measurement error is not related to our tracking method, but to the sensor. Still, we believe it would be possible to correct this instrument error by estimating the measurement area normal. For angles smaller than 30°, the error does not impact the measure, but at larger angles it does.

Processing times are given in Tables 2 and 4. They are linear with the number of samples used. The choice of 100 samples is a good compromise. Using the Kalman filter as well slows processing by about 2%. Better performances could be attained with a compiled language such as C++.

5 Conclusion

This paper presented a methodology to estimate the body temperature of a moving animal or of a human in a laboratory setting from thermographic camera images. The method begins

with the manual selection of a region of interest (ROI) on the subject, preferably exposed skin. The ROI is then automatically tracked with a particle filter that uses edge and intensity features as observation models. The average of the warmest pixels inside the tracked region is considered to be a temperature sample. We have observed that this approach increases resistance to small tracking errors in the particle filter. The temperature samples are then processed with a Kalman filter to reduce camera measurement noise.

The method was applied to six videos, five of Sprague-Dawley rats and one of a human. In order to validate the method, we manually processed randomly selected frames and compared the temperatures measured automatically to those measured manually. We found that the RMS error on the temperature measurements was usually less than $0.1°C$, and as low as $0.03°C$. Large errors were caused by severe occlusions, quick target motion, and possibly the change in orientation between the tracked surface and the camera sensor. The main advantage of the method is that it is non-invasive, resulting in no interference with the subject during an experimental protocol where the subject's core temperature fluctuations must be monitored.

Acknowledgements The authors thank the Canada Foundation for Innovation (CFI) for their grant for our measurement equipment. We also thank Cynthia Orman for revising the paper and the anonymous reviewers for providing helpful comments.

References

1. Acharya, R., Ng, E.Y.K., Yee, G.C., Hua, T.J., Kagathi, M.: Analysis of normal human eye with different age groups using infrared images. Journal of Medical Systems (33), 207–213 (2009)
2. Ahlenius, S., Oprica, M., Eriksson, C., Winblad, B., Schultzberg, M.: Effects of kainic acid on rat body temperature: unmasking by dizocilpine. Neuropharmacology **43**(1), 28–35 (July 2002)

3. Amalu, W.: Nondestructive testing of the human breast: the validity of dynamic stress testing in medical infrared breast imaging. In: Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE, vol. 2, pp. 1174–1177 (2004)
4. Camenzind, M., Weder, M., Rossi, R., Kowtsch, C.: Remote sensing infrared thermography for mass-screening at airports and public events: Study to evaluate the mobile use of infrared cameras to identify persons with elevated body temperature and their use for mass screening. Tech. Rep. 204991, EMPA Materials Science and Technology (2006)
5. Caniou, J.: Passive Infrared Detection, Theory and Application. Kluwer Academic Publishers (1999)
6. Carmant, L.: Mechanisms that might underlie progression of the epilepsies and how to potentially alter them. Advances in neurology **97**, 305–314 (2006)
7. Chiang, M., Lin, P., Lin, L., Chiou, H., Chien, C., Chu, S., Chiu, W.: Mass screening of suspected febrile patients with remote-sensing infrared thermography: Alarm temperature and optimal distance. Journal of the Formosan Medical Association **107**(12), 937–944 (2008)
8. Duda, R., Hart, P.: Pattern Classification and Scene Analysis. John Wiley and Sons (1973)
9. Forsyth, D.A., Ponce, J.: Computer Vision: A Modern Approach. Prentice Hall (2002)
10. Guntheroth, W., Spiers, P.: Thermal stress in sudden infant death: is there an ambiguity with the rebreathing hypothesis? Pediatrics **107**(Series D), 693–698 (2001)
11. Isard, M., Blake, A.: Condensation – conditional density propagation for visual tracking. International Journal of Computer Vision **29**(1), 5–28 (1998)
12. Jones, B.: A reappraisal of the use of infrared thermal image analysis in medicine. Medical Imaging, IEEE Transactions on **17**(6), 1019–1027 (1998)
13. Kalman, R.E.: A new approach to linear filtering and prediction problems. Transactions of the ASME–Journal of Basic Engineering **82**(Series D), 35–45 (1960)
14. Levesque, M., Langlois, J.P., Lema, P., Courtemanche, R., Bilodeau, G.A., Carmant, L.: Synchronized gamma oscillations (30-50hz) in the amygdalo-hippocampal network in relation with seizure propagation and severity. Neurobiology of Disease **35**(2), 209–218 (August 2009)
15. McDougali, A., Salter, D.: Thermography of the nose and ear in relation to the skin lesions of lepromatous leprosy, tuberculosis, leishmaniasis, and lupus pernio. Journal of Investigative Dermatology (68), 16–22 (1977)
16. Merla, A., Mattei, P.A., Di Donato, L., Romani, G.L.: Thermal imaging of cutaneous temperature modifications in runners during graded exercise. Annals of Biomedical Engineering **38**(1), 158 – 163 (2010)

17. Morgan, P., Smyth, J., Tullo, A., Efron, N.: Potential applications of ocular thermography. Optometry and Vision Science **70**(7), 568–576 (1993)

18. Ng, E.K., Fok, S., Peh, Y., Ng, F., Sim, L.: Computerized detection of breast cancer with artificial intelligence and thermograms. Journal of Medical Engineering and Technology **26**(4), 152–157 (2002)

19. Parsi, K., Kossard, F.: Thermosensitive lichen amyloidosis. International Journal of Dermatology **43**(12), 925–928 (2004)

20. Russell, M., Vink, R.: Increased facial temperature as an early warning in sudden infant death syndrome. Medical Hypotheses **57**, 61–63 (2001)

21. Shastri, D., Merla, A., Tsiamyrtzis, P., Pavlidis, I.: Imaging facial signs of neurophysiological responses. Biomedical Engineering, IEEE Transactions on **56**(2), 477–484 (2009)

22. Sunderam, S., Osorio, I.: Mesial temporal lobe seizures may activate thermoregulatory mechanisms in humans: an infrared study of facial temperature. Epilepsy and Behavior **49**(4), 399–406 (August 2003)

23. Torabi, A., Bilodeau, G.A., Levesque, M., Langlois, J.M., Lema, P., Carmant, L.: Measuring an animal body temperature in thermographic video using particle filter tracking. In: ISVC '08: Proceedings of the 4th International Symposium on Advances in Visual Computing, pp. 1081–1091. Springer-Verlag, Berlin, Heidelberg (2008)