The use of infrared thermography for the dynamic measurement of skin temperature of moving athletes during competition; methodological issues

Polly E Aylwin¹, Sebastien Racinais², Stéphane Bermon³,4, Alex Lloyd¹, Simon Hodder¹ and George Havenith¹,*

¹ Environmental Ergonomics Research Centre, Loughborough University, United Kingdom
² Aspetar Orthopaedic and Sports Medicine Hospital, Qatar
³ World Athletics, Health and Science Department, Principality of Monaco, Europe
⁴ LAMHESS, Université Côte d’Azur, France

* Author to whom any correspondence should be addressed.
E-mail: g.havenith@lboro.ac.uk

Keywords: infrared thermography, skin temperature, long wave infrared, medium wave infrared, competition

Abstract

Objective. To investigate the use of infrared thermography (IRT) for skin temperature measurement of moving athletes during competition and its sensitivity to factors that are traditionally standardised.

Approach. Thermograms were collected for 18 female athletes during the 20 km racewalk at the 2019 World Athletics Championships, with a medium-wave, cooled indium antimonide medium wave infrared band (MWIR) and a long-wave, uncooled microbolometer longwave infrared band (LWIR) infrared camera.

Main results. The MWIR provided greater clarity images of motion due to a shorter exposure and response time and produced a higher percentage of acceptable images. Analysing acceptable images only, the LWIR and WMIR produced good levels of agreement, with a bias of $-0.1 \pm 0.6 \, ^\circ\text{C}$ in mean skin temperature for the LWIR. As the surface area of an ROI was reduced, the measured temperature became less representative of the whole ROI. Compared to measuring the whole area ROI, a single central pixel produced a bias of $0.3 \pm 0.3 \, ^\circ\text{C}$ (MWIR) and $0.1 \pm 0.4 \, ^\circ\text{C}$ (LWIR) whilst using the maximum and minimum temperature pixels resulted in deviations of $1.3 \pm 0.4 \, ^\circ\text{C}$ and $-1.1 \pm 0.3 \, ^\circ\text{C}$ (MWIR) and $1.2 \pm 0.3 \, ^\circ\text{C}$ and $-1.3 \pm 0.4 \, ^\circ\text{C}$ (LWIR). The sensitivity to air and reflected temperatures was lower for the LWIR camera, due to the higher emissivity of skin in its wavelength.

Significance. IRT provides an appropriate tool for the measurement of skin temperature during real-world competition and critically during athlete motion. The cheaper LWIR camera provides a feasible alternative to the MWIR in low rate of motion scenarios, with comparable precision and sensitivity to analysis. However, the LWIR is limited when higher speeds prevent the accurate measurement and ability to capture motion.

Introduction

Body temperature is often characterised in two parts; the core and the skin. While core temperature is associated with risk of heat illness, skin temperature has been shown to be strongly associated with performance (Schlader et al 2011, Sawka et al 2012, Foster et al 2021). Skin temperature is a function of cutaneous blood flow, blood temperature, and dry and evaporative heat loss from the skin’s surface. Sweat secretion and subsequent evaporation can offset increases in skin temperature associated with increased hot, cutaneous blood flow. Even at air temperatures above 35 °C, evaporative heat loss can lower skin temperature below that of air temperature (Shapiro et al 1980). However, any impediment to evaporative heat loss, will raise skin temperature (Shapiro et al 1982).
Skin temperature is also a direct driver of autonomic thermoregulation, specifically blood flow regulation. Skin temperature acts directly on cutaneous blood vessels (Johnson 2010), contributing to thermoregulatory reflex responses. Whilst thermoneutral, tonic vasoconstriction is present and active vasoconstriction stimulated when skin temperature drops. Conversely, when skin temperature increases beyond 30 °C, vasoconstrictor tone is withdrawn and vasodilation stimulated (Hodges et al 2008, Wingo et al 2009); with peak skin blood flows occurring at 37 °C skin temperature. Thus, skin temperature is an important indicator of the human thermophysiological response and its measurement is therefore often applied in research fields (Coull et al 2021), clinical diagnosis (Fernández-Cuevas et al 2017), sports physiology (Perpetuini et al 2021), environmental building ergonomics (Ghahramani et al 2016) and clothing design (Raccuglia et al 2019).

Contact sensors
The most commonly used methods for measuring skin temperature are contact methods, using thermocouples or thermistors. Contact methods are non-invasive with good levels of accuracy, sensitivity and reliability (Harper-Smith et al 2010), along with a short response time and wide temperature measurement range (de Andrade Fernandes et al 2014). The disadvantages of these are high-cost data loggers and their wired nature, impeding the useability with non-stationary participants. A wireless alternative, iButtons, provide a cheaper more feasible device in a wider range of scenarios (van Marken Lichtenbelt et al 2006, Harper-Smith et al 2010). However, the embedded data logger means data cannot be accessed in real-time.

All contact methods suffer the limitation that the device itself and/or the attachment method applies insulation and pressure to the skin’s surface; creating possible interference of heat exchange, changes in the local skin temperature and alterations in evaporation (MacRae et al 2018a, 2018b), that must be accounted for (Buono et al 2007, Tyler 2011). Furthermore, the measurements can be influenced by ambient conditions and surrounding microclimate (MacRae et al 2018a, 2018b, Havenith and Lloyd 2020) which in turn altered by the type of tape and clothing coverage (Buono and Ulrich 1998, Tyler 2011, Psikuta et al 2014, MacRae et al 2018a, 2018b). Moreover, for long duration exercise, sweat and body movements can cause the device to detach (Buono et al 2007). Additionally, as a measurement of whole body skin temperature, they cover only a very small body surface area (BSA), as the sensor is small (Harper-Smith et al 2010) and cover typically 4–14 body sites (Ramanathan 1964, ISO 2004). This provides a very localised measurement, which may not be representative of the whole body, due to strong intra-regional variations (Livingstone et al 1987, Coull et al 2021).

Infrared thermography (IRT)
To overcome certain limitations of contact methods, an alternative tool is IRT. IRT has been used in thermophysiological research (Clark et al 1977, Costello et al 2012, Fournier et al 2013, Gerrett et al 2015, Tanda 2015, 2018, Ganse and Degens 2020, Maley et al 2020) and clinical practice (Clark and Stothers 1980, Ring and Ammer 2012, Fernández-Cuevas et al 2017) for the assessment of skin temperature. This provides a safe, non-invasive and potentially moderate cost technique, with the additional benefit of being non-contact and portable (Hildebrandt et al 2012, de Andrade Fernandes et al 2014). Infrared cameras process emitted thermal energy from the object studied. Any surface with a temperature above absolute zero emits thermal energy which falls within the infrared portion of the electromagnetic spectrum. An infrared camera receives three types of radiation; that which is directly emitted by the object, emissions from ambient sources that are reflected by the object and emission from the atmosphere between camera and object. In most situations the emission from the object is the most important parameter however, control of all the other factors is important (Gomes-moreira et al 2017). In field testing where the external radiation and atmospheric conditions are less well controlled, these parameters are of even greater relevance than in lab studies.

IR detector types
IR cameras can have quantum or thermal detectors, which absorb the emitted infrared radiation. Quantum detectors, image in the medium wave infrared band (MWIR), from 2 to 5.5 μm; whilst thermal detectors typically image in the longwave infrared band (LWIR) from 7 to 14 μm. Thermal detectors such as microbolometers absorb radiant heat, which heats a pixel within a sensor matrix causing a change in electrical resistance that can be measured and displayed as an image (Rogalski 2010, pg 104). Thermal detectors do not require cooling and cost less than quantum detectors (Rogalski 2010, pg 31). However, as cooled quantum detectors do not have to change temperature, instead operating as a semiconductor infrared photon detector, they have a much faster response time (Rogalski 2010, pgs 31 and 176), making it possible to image motion and/or measure very rapid temperature changes (note: the response time is the time constant for the output value to respond to a step change in the object temperature i.e. at the response time, the output value has reached 63% of its final steady state value, thereby requiring 5 times the response time to get to within 1% of the final value) (FLIR 2015). They
also have quick exposure times (time to capture a single frame) and frame rates (how rapidly frames can be imaged and how frequently), further enhancing their ability to capture motion.

IRT for skin temperature measurement
IRT provides an accurate (Roy et al 2006), reliable (Selfe et al 2006) and highly sensitive (±0.1 °C) method. In line with previous definitions (Fernández-Cuevas et al 2015) we define accurate, as the ability of a measurement to provide the ‘true’ value, reliability as the ability of a measurement to produce consistent results across multiple measurements, when the true value is stable and sensitivity as the smallest change/difference that the measurement can detect. The accuracy of IR cameras is typically ±1 to 2 °C, which is poor compared to contact methods, that are typically ±0.1–0.5 °C, however this can be improved by including a reference temperature within images. IRT has previously been found to deviate from thermistors by 1.80 ± 1.16 °C (Maley et al 2020) and 1.88 ± 1.87 °C in the recovery from exercise (Bach et al 2015b). However, both studies should be approached with caution (Havenith and Lloyd 2020, counterpoint to Maley et al 2020). Firstly, they utilise thermistor measurements as their reference skin temperature, without accounting for the inaccuracies these can provide, as discussed above. Secondly, Bach et al (2015a) do not reference the size of ROI or pixel count and Maley et al (2020) analysed a single pixel. It may be the case that these do not produce an accurate weighted mean skin temperature (Mean $T_{sk}$), due to local variation in skin temperature caused by perforator vessels or vasculature (Merla et al 2010, Chudecka and Lubkowska 2012). Finally, neither study references using a black body calibration, which can improve accuracy (Gomes-moreira et al 2017).

Implementing IRT
Standardisation of the protocol for thermography, has been appreciated for some time; in 1997 the European Association of Thermography laid out a standard for medical applications (Clark and de Calcina-Goff 1997). Factors such as standardising the environmental conditions of the examination room, and the analysis methods were proposed. This has been followed by further work for clinical assessment (Ring and Ammer 2015) and physiological research in the form of a Delphi study derived checklist (Gomes-moreira et al 2017), however the factors on the checklist were chosen for laboratory based research. Factors noted as requiring standardisation include: the use a reflective black reference plate calibration (or an IR calibrator unit), use of the correct input parameters (emissivity, reflected temperature, air temperature, relative humidity and object distance), stable environmental conditions and consistent set-up for the thermograph. The environmental conditions are consistently noted as needing to be stable and controlled, as variability can impact results (Bach et al 2015a). Participants should be static, as otherwise unacceptable blurring can occur, as seen previously in the work of Gans and Degens (2020). Furthermore, they display the requirement for good temperature contrast between the person and the background to avoid loss of distinction, in the colour palette selected (Tanda 2018). Finally, if the distance of the object is inconsistent the variable number of pixels can lead to inaccuracies (Ring and Ammer 2015). However, as these cannot be controlled in field-testing it is important to assess the contribution they make to the accuracy of results with moving targets. Some studies have tried to implement IRT during exercise, but often during breaks in activity (de Andrade Fernandes et al 2014) or without imaging fast moving body areas (Merla et al 2010, Tanda 2018, Hadzić et al 2019). These however, do not provide representative measures of whole-body skin temperature and cannot be utilised during competition where athletes cannot be stopped to be imaged.

The input parameters required by the post-processing software include emissivity of the object, reflected temperature (a measurement of the infrared collected by the camera that originated from other sources and reflected off the object in question), air temperature, relative humidity and distance to the object. Skin emissivity was first appreciated by Hardy and Muschenheim (1934) and well understood to be high and constant, not too far from a black body. The most referenced value is 0.98, from Steketee (1973), also agreed upon by 96% of experts in a Delphi study (Gomes-moreira et al 2017), though that paper assumed application of LWIR only. However, there is still some debate, as others show it to be between 0.95 and 0.99 (Hardy and Muschenheim 1934, Jacques et al 1955, Togawa 1989, Sanchez-Marin et al 2009). Notably, the magnitude of error in measurement of skin temperature for different emissivities and input parameters have not been reported.

The potentially subjective nature of analysis of IRT, has brought into question the best method of determining ROIs, to maximise reliability. Methods have been proposed for automating or reducing possible observer bias, such as selecting the hottest pixel and surrounding 24 (Ludwig et al 2014) or the hottest 5 pixels (Formenti et al 2017). These automated methods are supported in standardisation protocols (Fernández-Cuevas et al 2015). However, there is no clear consensus across all literature; perhaps because beyond improving reliability, these have significant limitations. These standardised methods commonly analyse a smaller surface area of the skin than traditional methods and create a bias towards analysing the hottest areas. The loss of
information from the thermogram, negates the benefit of being able to analyse a large BSA; particularly important where ‘mottling’, due to vasculature or perforator vessels is present. However, whether this makes a meaningful impact to measurements of skin temperature, remains to be investigated.

The standardisation of IRT has led to its use primarily in well controlled laboratory studies, with stable environmental conditions, static participants and strict control. Despite the numerous benefits for field-testing, its use in this setting has been more limited, and research to date has not appreciated the impact of the limited standardisation possible. Therefore, the aims of this study were to:

(1) Establish whether IRT can be used to collect skin temperature data of exercising athletes, during competition, using MWIR and LWIR cameras

a. Do the results of the MWIR and LWIR agree, and can they be used interchangeably?

b. Are the response time and capture rate of each camera sufficient to image moving athletes, without inhibitive blurring?

c. What issues prevent a thermogram or specific body site from being analysed?

i. Is this altered by the colour palette selected for analysis?

d. What is the extent of missing data and how sensitive is the mean $T_{sk}$ calculation, to individual and multiple missing ROIs?

(2) Determine the impact of different size ROI selection methods.

(3) Investigate the influence of factors that cannot be controlled in field-based competition.

a. How sensitive is each camera to alterations in input parameters; emissivity, reflected temperature, air temperature, relative humidity and object distance?

b. How do results differ for static measurements versus during motion?

Methods and results

Participants and procedure

As part of a larger study, thermal videos were collected during road races at the 2019 IAAF World Athletics Championships, Doha. A subset of those thermal videos was investigated here prior to analysis of the full data set. The data utilised was that of 18 females competing in the 20 km racewalk. This paper focussed on the investigation of the infrared methodology and a comparison of the two camera types dictated primarily by the physics of the methodology not the physiology of the human investigated, and thus we concluded that using a single sex in the study was acceptable. The project was approved by the Anti-Doping Laboratory Qatar ethics committee (E2019000302). All procedures complied with the Declaration of Helsinki and written informed consent was obtained.

Environmental conditions were measured every 30 min using a heat stress meter (Kestrel 4400, Boothwyn, PA) on a tripod ∼1.5 m above the floor, ∼4 m from the course. The mean conditions over the duration of the race were 31.6 ± 0.8 °C air temperature, 76 ± 3% relative humidity, 0.3 ± 0.4 m·s⁻¹ air velocity and 29.0 ± 0.5 °C WBG. The race took place on a flat track of 1 km loops, between the hours of 23:30 and 02:00, during darkness when air temperatures are lower and solar radiation removed. Although body temperatures fluctuate with circadian rhythm, both core and skin temperature responses during exercise, have been seen not to be altered by the time of day of the test (Ravanelli and Jay 2021). Both IR and a normal camera were placed on a U-turn in the track, to allow measurement of the anterior (at a distance of ∼9 m), lateral (at a distance of ∼7 m) and posterior (at a distance of ∼9 m) aspects of athletes. Two cameras were used to collect infrared thermal video and a third camera collected normal video for participant identification. The MWIR camera refers to a cooled indium antimonide camera (FLIR A6750sc MWIR, FLIR Systems, West Malling, Kent; 640 * 512 Focal Plane Array; 3–5 μm spectral range; 20 mK noise equivalent temperature differences (NETD); 15–30 Hz sampling rate; f/2.5; accuracy ±2 °C). The LWIR camera refers to an uncooled microbolometer camera (FLIR T1030sc; 1024 * 768 focal plane array; 7.5–14 μm spectral range; 20 mK at 30 °C NETD; 30 Hz sampling rate; f/1.15; accuracy ±1 °C). The two cameras were stabilised for a period of ∼45 min prior to the start of the race. Two 12*12 cm² plates were mounted and positioned to the side of the track; one crumpled aluminium foil coated for reflected temperature and one matte black with a thermistor to act as a black-body calibration reference. ROIs were manually selected in FLIR software (FLIR ResearchIR Version 4.40. 11.35), all by the same researcher. ROI data were collected at kilometers 1, 5, 10, 15 and 20 unless otherwise stated. Statistical analysis included Bland–
Altman and T-tests for comparisons of ROI sizes and the three imaged perspectives. The alpha level was set to $p < 0.05$.

**Analysis methods and findings**

As per the research aims the following sections lay out the factors investigated, detailing the method and findings:

**Camera type**

As discussed, MWIR cooled quantum detector cameras have a shorter exposure time, response time and allow higher sampling rates than uncooled LWIR microbolometer cameras. The movement of athletes through the field of view was therefore expected to be more problematic in the LWIR camera than the MWIR. Blurring of images is an issue for accuracy of measurements and reliability of ROI placement. To establish whether the LWIR camera was able to provide accurate measurement in this application the images and results of the two cameras were compared for the 18 athletes.

**Identifying images satisfactory for analysis**

In laboratory studies, participants are typically asked to stand in a fixed (anatomical) position, in a fixed location (Fournet et al. 2013, Gerrett et al. 2015). However, utilising IRT for exercising athletes, during competition means that the exact thermogram obtained cannot be standardised. Potential issues include blurring, obstructions by other athletes or external cooling devices, multiple athletes passing at once meaning the camera cannot track them to obtain an anterior and posterior image and unclear definition of the edges of ROIs due to similar skin and surrounding temperatures (in this case only $\sim 0.5 \, ^\circ C$ different).

Figures 1(A)–(C), display example thermograms from the MWIR camera that were considered ideal for analysis i.e. body parts are clear and unobstructed by any surrounding object, the edges can be easily identified and the whole ROI area can be collected. As can be seen for the lateral perspective (figure 1(B)) all the chosen ROIs (listed in table 1) could be analysed from a single thermogram. This was frequently possible with the lateral measurements, but less frequently in the anterior and posterior as one call or forearm was not perpendicular to the camera at the same time as the other (figures 1(A) and (C)). Additionally, in the anterior perspective arms often crossed over the body, covering the torso (figure 1(A)). These factors meant that anterior and posterior measurements were more commonly taken over two or three sequential frames. Ideal images from the LWIR camera (figures 1(D)–(F)) were less frequent and more frequently only had one or two ROIs identifiable per image, due to high levels of blurring. The posterior perspective provided the greatest clarity thermograms, whereas blurring was a greater issue in the anterior lower limbs and particularly all areas in the lateral perspective (figure 2). The degree of blurring ranged from slight blurring of the edges, through to a transparent appearance of the ROI.

**Causes of blurring**

Blurring is an important issue in moving objects; determined by the speed of the moving person across the image pixels and several camera properties (time constant and integration time, image capture time and method). To get a valid temperature measurement, the object must remain on the sensor matrix pixels for a sufficient time. A microbolometer pixel absorbs the IR energy from the source over time, for which we need to consider the time constant for the pixel to reach a stable temperature. For the LWIR, the time constant (response time; to reach 63% of the final value) is typically around 10 milliseconds, thus requiring the object to remain on the pixel for around 50 ms to reach a stable (within 1%) readout of the actual temperature. If the image stays on the pixel for less, due to movement, the measurement averages over a bigger area of the object, and in the worst case the pixel may for part of the time be seeing the background rather than the object. For the MWIR quantum detector camera, given the different technology, this is based on the time taken to collect sufficient photons: the integration time. This is between 5–6 microseconds and a few milliseconds (FLIR Systems Inc. 2015 and Raphael Danjoux, FLIR Systems Inc., personal communication) with 0.1 ms taken as an approximation for following calculations. This integration time is also the time required to get to the actual value.

An estimate of the movement artefact for the lower leg area close to the foot can be made as follows:

\[
\text{Pixels moved per second} = \text{Walking steps per second} \times \text{no. of pixels moved per step}
\]  \hspace{2cm} (1)  

\[
\text{Pixels moved per exposure} = \text{pixels moved per second} \times \text{exposure time (seconds)}
\]  \hspace{2cm} (2)

World class athletes run at speeds of for example, $\sim 9 \, \text{m} \cdot \text{s}^{-1}$ for a 400 m sprint, $\sim 7 \, \text{m} \cdot \text{s}^{-1}$ for a 1500 m and in this case the mean movement speed of athletes was 3.56 $\, \text{m} \cdot \text{s}^{-1}$, if we assume this to be $\sim 3$ steps (one step being the action of a single foot leaving the ground to being placed back down) per second (each step $\sim 1$ m in
distance), in the images a single leg in lateral view is therefore moving at approximately 630 pixels per second, on the LWIR camera. Investigating the fastest moving body area, the legs, this means during an exposure time of 30 ms, the leg moves over 19 pixels. This firstly shows that the leg image is not long enough on the same pixel to let the pixel measure the temperature of that specific leg area. Further, given that the leg in this case was only 12 pixels wide, this means that in the exposure time of 30 ms, 7 pixels are at one point in the exposure actually

| Table 1. Weighting for each ROI, for calculation of weighted mean exposed skin temperature ($\text{mean } T_{sk}$) for each aspect. |
|---|---|---|
| Anterior | Lateral | Posterior |
| Face | 0.081 | 0.043 | — |
| Neck | 0.037 | — | — |
| Torso | 0.157 | 0.043 | 0.098 |
| Hands | — | 0.025 | — |
| Shoulders | 0.157 | 0.039 | 0.103 |
| Upper arm | 0.073 | 0.092 | 0.106 |
| Forearm | 0.038 | 0.073 | 0.063 |
| Thigh | 0.349 | 0.399 | 0.388 |
| Calf | 0.202 | 0.253 | 0.243 |

Figure 1. Thermograms taken on the cooled MWIR camera (A)–(C), that are considered ideal for analysis, for the anterior, lateral and posterior perspectives. Thermograms from the uncooled LWIR camera (D)–(F), that show certain ROIs as ideal for analysis, with other ROIs unable to be clearly defined or blurred (lower limbs in D and E).

Physiol. Meas. 42 (2021) 084004 P E Aylwin et al
partially background measurements and not on the leg for the whole exposure. When we compare this to the MWIR camera, running at the same pace a single leg moves 700 pixels per second (more pixels than the LWIR because the field of view is narrower). This equates to moving 0.07 pixels in a single capture of 0.1 ms. This shows that the image is static enough to achieve a sharp image of the object on each pixel within the integration time. With the leg 14 pixels wide in the image, only 1 pixel on the edge of the leg is including background temperatures due to movement, and this edge pixel is in any case not included in the analysis. The influence of the movement artefact is also dictated by how the data for each image frame matrix is read. The MWIR camera, reads the temperature of all the pixels at the same moment in time, like a ‘snapshot’. Thus, as the temperature over a whole ROI is being recorded at the same moment in time, in addition to the small movement artefact, means that in a single frame each reading of a single pixel corresponds to a fixed area of the ROI throughout. However, the LWIR camera reads each matrix row by row from top to bottom of the image in 30 ms (rolling shutter). Therefore, the pixels of an ROI at the bottom of the image are read later than those at the top, causing distortion of the moving image. In extreme cases, a moving straight leg could become curved. In our data this could not be separated from the general movement artefact, but its impact would be largest when the moving item is spread over a large part of the image array.

Blurring can also make distinguishing the edge of the ROI more difficult, critical for avoiding an ‘edge’ effect. Difficultly was caused by the similar surrounding temperatures (figure 2), particularly at the start of races where athletes were grouped together, as a mass start event and body parts overlapped (figure 3). This issue was improved by using colour palettes that make similar temperatures more distinctive or by expanding the temperature scale (figure 3). Further obstruction was caused by clothing: all athletes wore at least a cropped top, knicker style shorts and ankle socks and trainers, covering ~31 % of BSA. However, extra clothing, such as caps, long tops, longer shorts or socks, in some cases covered up to an additional ~30% BSA. Some individuals carried external cooling devices, which again impeded the ability to analyse certain ROIs; this was particularly the case for the neck and shoulders from the anterior perspective and the hands. Obstructions had the least impact upon
the posterior view as cooling devices did not cover the back and athletes tended to move apart as they exited the corner of the track. The mean difference in measurement between the two cameras, using only good quality images, was 0.1 ± 0.6 °C, with the MWIR having the slightly higher mean value. This relationship was stable across absolute skin temperatures and ROIs.

Colour palette
Image processing software, provides a selection of colour palettes that images can be viewed in. Most fall into two categories, those where the saturation and hue change in a linear fashion with temperature (e.g. fusion or greyscale). And those where the colour is not on a continuous scale, instead temperature changes are shown with a discrete scale (medical). Selecting the appropriate colour palette can improve object clarity and the ease of identifying body parts. Figure 3 shows the same frame displayed in three different colour palettes; the medical palette was useful where the edge of the athlete needed distinguishing from the surroundings, of a similar temperature. This is useful in this context where the analysis requires that a margin is left around the edge of an ROI, but the air temperature and skin temperature are similar. However, it was less useful where athletes were grouped together, producing high variability across the image. For visualisation of results, a colour palette such as fusion offers clearly identifiable colour change, with the gradient being intuitive for the reader. Greyscale provides less intuitive visualisation of overall results, but it was found useful for distinguishing subtle differences if required for analysis, such as the similar skin and air temperature in the hot climate of the race.

Surface area of the ROI
One of the potential major benefits of IRT is that skin temperature can be measured over a much larger surface area than is possible with contact methods (the mean nude surface area analysed for the athletes was 0.98 m², this was 64% of the mean total BSA of the 18 athletes (1.52 m²)), thereby accounting for local variation in skin temperature (e.g. mottling in the legs in figure 3, in the fusion palette). To assess whether the surface area size of the ROI influences the results, measurements were taken with a range of ROI sizes: taking the whole-body part (examples shown in figure 1), 200 pixels, 100 pixels, a straight line through the middle of the segment and a single pixel; all covering the centre point of the ROI. Additionally, from the full area ROI the hottest (maximum) and the coolest (minimum) pixel were compared. This was performed on two ROIs per participant on the 20th km, on both MWIR and LWIR cameras.
The bias from each of the smaller ROI areas are shown in figure 4 for the MWIR camera and figure 5 for the LWIR. For the MWIR camera, as the area of measurement was reduced a small positive bias ($0.2^\circ$C–$0.3^\circ$C, $p < 0.05$) compared to measuring the whole area developed which became slightly greater and showed greater variation (SD up to $0.3^\circ$C) the fewer pixels used, suggesting the more central pixels were typically the warmer ones. The LWIR camera also showed a small bias ($0.1^\circ$C, $p < 0.05$), positive for the reduced pixel counts, negative for the line, but the variability in the measurement increased substantially up to $0.4^\circ$C. Taking the maximum or minimum temperature from a body site showed the substantial variability of temperature within an ROI and resulted in large deviations from the actual mean of up to $\pm 1.3 \pm 0.4^\circ$C ($p < 0.05$) on both cameras, showing the potential risk of taking single pixels. While the mean of the group was not so different, the large variation across the ROI shows the potential risk for individual errors.

Additionally, the regional variation between the anterior, lateral and posterior view was investigated. ROI measurements obtained from two or more perspectives of the anterior, lateral and posterior were compared for all 18 participants at kilometers 5, 10, 15, 20, giving 412 data points. The mean difference in measurement was $0.4 \pm 0.6^\circ$C higher for anterior versus lateral, $0.2 \pm 1.0^\circ$C lower for anterior versus posterior ($p < 0.05$ for both) and a $0.02 \pm 3.0^\circ$C difference ($p > 0.05$) between lateral and posterior, shown in figure 6. Note the total
of the three differences do not cancel out, because the measurements of each are not taken at exactly the same moment in time, in a single frame.

**Mean skin temperature—impact of missing data points**

In addition to the ROI selection influencing skin temperature, it was hypothesised that missing individual body areas in the images for a specific lap would alter the weighted mean $T_{sk}$ value. Missing ROIs are inevitable in field-based conditions. For in-race measurements to calculate a weighted mean $T_{sk}$, ROIs were weighted, separately for each aspect, based on the mean number of pixels contained in the ROI (table 1). As the distance from the camera varied between aspects (anterior, lateral, posterior, shown in figure 1) the pixels represent different surface areas in each aspect. Therefore, the weightings were derived by converting pixel counts into a surface area measurement. To achieve this a known area (the athletes height) was measured in pixels for each perspective. This was converted to the number of pixels within a 1 m$^2$ area, which was used to calculate the surface area of each pixel, depending on the perspective of the athletes. The average number of pixels of ROIs for each body part were collected and the surface these covered calculated. The total surface area covered for the

![Bland–Altman plots comparing different ROI selections, to analyse skin temperature, from the LWIR camera.](image)

**Figure 5.** Bland–Altman plots comparing different ROI selections, to analyse skin temperature, from the LWIR camera. The X axis displays the average of the skin temperatures measured by the full area ROI and the alternative ROI selection and the Y axis displays the difference (alternative ROI—full area ROI) between the two. The alternative ROIs were 200 pixels, 100 pixels, central line, single pixel, the maximum temperature of a single pixel recorded in the full ROI area ($T_{max}$) and the minimum single pixel in the full ROI ($T_{min}$). The mean bias $±$ standard deviation of each alternative ROI is displayed. $T_{sk} = 23.0 °C; T_{air} = 30.5 °C; RH = 78%$; distance = 7–9 m; $\varepsilon = 0.98$. Each symbol represents an individual ROI, please note, where symbols appear darker this is due to overlapping data points. All were significantly ($p < 0.05$) different from the full area ROI, except the line.
anterior, lateral and posterior were then individually calculated and converted into a weighting, creating equation (3). For the pre and post measurements, the lateral weighting was redistributed equally to the anterior and lateral measurements.

\[ Weighted \text{ mean } Tsk (°C) = \frac{0.32 \text{ Anterior} + 0.38 \text{ Lateral} + 0.30 \text{ Posterior}}{3}, \]  

where \( Tsk \) = skin temperature.

To assess the impact of missing ROIs on mean exposed \( Tsk \), sites were removed individually and in combinations, from the calculation of mean skin temperature and compared with outcomes for the full data set. Where a site was removed the weighting of the missing site/s was distributed across the other sites that were included. Table 2 shows that the largest impact of missing ROIs on the mean \( Tsk \) value came from the thighs, with the highest pixel count and therefore weighting in the mean \( Tsk \) calculation. As expected, multiple missing sites had a larger impact on the mean \( Tsk \) than each constituent part alone. It was common in the 18 athletes for more than one body site to be missing and therefore how the missing data is dealt with requires consideration. When exercising in cool environments, with potentially larger \( Tsk \) variation, the absolute impact of missing ROI’s may be larger.

**Sample size**

Subsequently, it was investigated how many data points were needed for the mean of participants to be representative of the whole sample i.e. how many participants needed to be included for the mean to reach a stable value. To investigate this, the mean value of all participants for a body site was compared with all 18 participants included in the mean and by then reducing the number of participants by one at a time. The order in which participants were removed was worked through systematically, so that the specific individuals removed varied. Thus, giving a number of outcomes for the same ROI, with different combinations of participants. The mean remained very stable, on average within \( \pm 0.2 °C \), as the number of participants was reduced until reaching 4 participants, where differences in the mean become \( \pm 0.3 °C \) or greater.
Table 2. The absolute mean $T_{sk}$ (delta from none missing) calculated with all possible ROIs and with ROIs missing, from all aspects, and when missing from specific aspects individually. Where an ROI is missing its weighting was redistributed across the available sites.

| Site/s missing from calculation | Mean $T_{sk}$ (°C) | Mean $T_{sk}$ (°C), with anterior missing | Mean $T_{sk}$ (°C), with lateral missing | Mean $T_{sk}$ (°C), with posterior missing |
|--------------------------------|-------------------|------------------------------------------|-----------------------------------------|------------------------------------------|
| None                           | 32.3              | 32.3                                     | 32.3                                    | 32.3                                     |
| Face                           | 32.2 (±0.1)       | 32.2 (±0.1)                              | 32.2 (±0.1)                             | —                                       |
| Neck                           | 32.3 (0.0)        | 32.3 (0.0)                               | —                                       | —                                       |
| Torso                          | 32.2 (±0.1)       | 32.2 (±0.1)                              | 32.2 (±0.1)                             | 32.3 (0.0)                              |
| Hands                          | 32.3 (0.0)        | —                                        | 32.3 (0.0)                              | —                                       |
| Shoulders                      | 32.2 (±0.1)       | 32.2 (±0.1)                              | 32.2 (±0.1)                             | 32.3 (0.0)                              |
| Upper arms                     | 32.2 (±0.1)       | 32.3 (0.0)                               | 32.3 (0.0)                              | 32.3 (0.0)                              |
| Forearms                       | 32.3 (0.0)        | 32.3 (0.0)                               | 32.3 (0.0)                              | 32.3 (0.0)                              |
| Thigh                          | 33.7 (0.5)        | 32.5 (0.3)                               | 32.4 (0.1)                              | 32.3 (0.0)                              |
| Calf                           | 32.2 (±0.1)       | 32.2 (±0.1)                              | 32.2 (±0.1)                             | 32.3 (0.0)                              |
| Hands and torso                | 31.9 (±0.4)       | —                                        | —                                       | —                                       |
| Calves and thighs              | 32.9 (0.6)        | —                                        | —                                       | —                                       |
| Upper arms, forearms and thighs| 33.1 (0.8)        | —                                        | —                                       | —                                       |
| Face, hands and forearms       | 32.3 (0.0)        | —                                        | —                                       | —                                       |
| Torso, shoulders, forearms and calves | 31.9 (0.4)       | —                                        | —                                       | —                                       |
| Upper arms, forearms and thighs| 33.1 (0.8)        | —                                        | —                                       | —                                       |

Sensitivity to input parameters

For accurate analysis, the following parameters are required in the post-processing software or the camera itself: object emissivity, object distance, reflected temperature, air temperature and relative humidity. Emissivity is a measure of how well an object emits infrared radiation in relation to its temperature compared to a black body; for analysis of skin temperature it is traditionally set between 0.95–0.98, most commonly 0.98 (Steketee 1973). However, the emissivity should not be assumed to fall within this range for all cameras. The MWIR camera measures infrared in the 3–5 μm wavelength band. At these wavelengths it had been suggested skin emissivity should be set at 0.91–0.92 (Gaussorgues 1994), therefore, for the analysis in this paper, emissivity was set at 0.91 for the MWIR camera, unless otherwise stated.

The reflected temperature ($T_{refl}$) is any radiation from surrounding objects that reflects off the object of interest, a value that can be obtained by the crumpled foil method; by placing a square coated with crumpled aluminium foil within the frame and taking a reading, with the emissivity set to 1. This is not well described in the literature as being required as part of the standardised thermal imaging procedure, with it being omitted by standardisation protocols from Fernández-Cuevas and Gomes-moreira et al (2017) and Ring and Ammer (2012). The sensitivity of the measurement outcomes to deviations in each parameter independently and combined were investigated and the results shown in table 3.

The interactive relationship of emissivity, reflected temperature ($T_{refl}$) and air temperature ($T_{air}$)

The greater the difference between $T_{refl}$ and the object temperature ($T_{object}$, in this case skin temperature) the greater the impact of changing emissivity. Conversely, the sensitivity to changing $T_{refl}$ was dependent on the emissivity. At an emissivity of 1, changing the $T_{refl}$ had no impact on $T_{object}$. But as emissivity lowers, the sensitivity to $T_{refl}$ increases linearly. The relationship between emissivity, $T_{air}$ and $T_{object}$ was the same linear pattern. However, $T_{object}$ sensitivity to $T_{air}$ is less than $T_{refl}$. The relationships were the same for both the MWIR and LWIR camera. The only difference being the LWIR camera was more sensitive to a deviation in $T_{air}$. Furthermore, the sensitivity differed depending on whether $T_{air}$ was higher or lower than $T_{object}$.

Sensitivity to relative humidity and distance dependent on $T_{air}$

The sensitivity to relative humidity was dependent on $T_{air}$ but not $T_{refl}$. However, at 0.91 and 0.98 emissivity, both cameras were unaffected by a deviation of ±5%. Given that the effect of relative humidity is related to the amount of vapour in the air (vapour pressure or g.m$^{-3}$), the impact would be different for identical relative humidity values at different temperatures. However, given the high vapour pressure taken in the calculation example showing minimal impact, the relative humidity deviation is of limited importance at the distances tested here. The deviation introduced by a deviation in distance was linear, across ±4 m from 7 m. The sensitivity to distance was independent of $T_{refl}$ but dependent on the difference between $T_{air}$ and $T_{object}$. When $T_{air} > T_{object}$ this is a negative error and when $T_{air} < T_{object}$ a positive error.
Table 3. The deviations in measured object temperature ($T_{\text{object}}$) introduced by a deviation in input parameter and the sensitivity of each camera to deviations in each input parameter, along with the interactions between each input parameter.

| Parameter | Error size tested | MWIR | LWIR |
|-----------|-------------------|------|------|
| $T_{\text{air}}$ at 0.91°C | ±4°C | Error in $T_{\text{object}}$: 0.80°C | Error in $T_{\text{object}}$: 0.161°C |
| $T_{\text{refl}}$ at 0.91°C | ±4°C | Sensitivity: 0.2°C/1°C error | Sensitivity: 0.040°C/1°C error |
| $T_{\text{air}}$ at 0.98°C | ±4°C | Error in $T_{\text{object}}$: 0.08°C | Error in $T_{\text{object}}$: 0.054°C |
| RH (at $T_{\text{air}}$ = $T_{\text{object}}$) | ±5% at 75% | Error in $T_{\text{object}}$: <0.1°C | Error in $T_{\text{object}}$: <0.1°C |
| Distance | ±2 at 7 m | Error in $T_{\text{object}}$: 0.009°C/1 m error | Error in $T_{\text{object}}$: 0.009°C/1 m error |

**MWIR**

- $T_{\text{refl}}$ at 0.91°C: ±4°C, ±0.07°C at 0.91°C
- $T_{\text{air}}$ at 0.98°C: ±4°C, ±4°C
- $T_{\text{air}}$ at 0.91°C: ±4°C
- $T_{\text{air}}$ at 0.98°C: ±4°C
- RH (at $T_{\text{air}}$ = $T_{\text{object}}$): ±5% at 75%
- Distance: ±2 at 7 m

**LWIR**

- $T_{\text{refl}}$ at 0.91°C: ±4°C, ±0.07°C at 0.91°C
- $T_{\text{air}}$ at 0.98°C: ±4°C, ±4°C
- $T_{\text{air}}$ at 0.91°C: ±4°C
- $T_{\text{air}}$ at 0.98°C: ±4°C
- RH (at $T_{\text{air}}$ = $T_{\text{object}}$): ±5% at 75%
- Distance: ±2 at 7 m
The impact of changing skin wettedness

With skin wetting, the emissivity of the skin changes; in the LW range, emissivity of pure water is around 0.97 (Robinson and Davies 1972) to 0.99 and virtually identical for salt water (+0.003) between 11 and 12 μm (Masuda et al 1988) changing minimally with viewing angles up to 30° (i.e. limited reflection). This would imply an emissivity change of a negligible 0.01 for the LWIR camera. In this wavelength range, IR penetration through the water is low (Hale and Querry 1973), indicating the temperature taken will be that of the water surface. For thermal balance calculations, it is the surface values of temperature and vapour pressure that are relevant, whether that is the skin when dry or the water surface when wet.

For the MWIR camera changing from skin to water could have a bigger effect, as the difference in emissivity is larger. Emissivity changes from 0.91 for skin to 0.97 for water, around 3–7 μm (Masuda et al 1988). However, in this range IR penetration through the water is much higher (Hale and Querry 1973) up to 1 mm, indicating that for thin water/sweat layers in this MW spectral range the skin ‘shines through’, i.e. the resultant emissivity lies between the skin and water value, reducing the impact of the wetting emissivity change on the outcome values. The latter is supported by the low bias between the MWIR and LWIR cameras observed (0.1 °C), where emissivity was not adjusted for wetting. With increasing temperature differences between the skin/water surface and the environment, the effect however can become larger, especially for the MWIR camera and viewing angles should not deviate more than 30°–40° from perpendicular.

Comparison of static and moving skin temperatures

Prior to and immediately following the racewalk, static images of each of the 18 athletes were taken from the anterior and the posterior perspectives, to assess how representative static data is for skin temperature during exercise. Post exercise thermograms were taken within fifty meters of the finish, before athletes entered the media area. This occurred only a couple of minutes after crossing the finish line. In this case, weighted mean skin temperature was calculated with the anterior, lateral and posterior during the race and the anterior and posterior only post-race. The mean $T_{sk}$ rose by a mean of 4.5 °C, from the final lap to post (figure 7). For comparison, as the pre- and post-static images are only of the anterior and lateral, if the mean $T_{sk}$ includes only the anterior and posterior during the race, the mean $T_{sk}$ post-race was 3.5 °C higher.

Discussion

This study aimed to investigate whether IRT could be used to collect exposed skin temperature data of moving athletes, during competition. Secondly, to investigate which factors the method is sensitive to if unstandardised. The main findings were good agreement in measurements of a long wave and a medium wave infrared camera for sites that could be analysed. However, due to the impact of motion on the usability of individual images for the LWIR, with difficulties capturing sharp images of the moving arms and legs and of the whole body in the lateral aspect, the quantum based MWIR camera performs much better for field measurements during exercise. Given the lower skin emissivity value in the IR range of the MWIR (Gaussorgues 1994) this is however more sensitive to deviations in environmental parameters, as well as to emissivity changes when the skin is wet. The small difference between cameras over all measurements suggests that these effects are limited for the conditions of the race studied. A reduction in the surface area of an ROI reduced its representativeness for the temperature of the ROI covering the whole area. Relevant differences in skin temperatures between the three aspects (anterior, lateral and posterior) were present as well as a meaningful impact of missing data; supporting analysing as large a surface area as possible. Both cameras were most sensitive to deviations in emissivity and air and reflected temperatures. IRT provides a portable, non-contact temperature measurement device that can be utilised to measure nude skin temperatures of athletes during motion. Appropriate use of IRT for this purpose will therefore allow field-based, measurements of skin temperature, without having to pause exercise to make measurements. This allows greater use during competitive events, furthering our understanding of human physiological responses to environmental conditions in real life scenarios.

Camera type

Both a cooled quantum detector MWIR and an uncooled microbolometer detecting LWIR camera, were able to image moving athletes (but with substantial motion artefact limitations for the LWIR) and provided good agreement in skin temperature readings. The MWIR, given its shorter response and exposure times provided more consistent quality and usability of images, allowing more ROIs to be analysed in a single frame, improving reliability of measurements and speed of analysis. Blurring/movement artefacts from the LWIR camera were particularly an issue for the faster moving legs and forearms in the frontal plane and the whole body whilst moving through the lateral field of view.
The appropriate application of IR to moving participants has been mixed, with some producing quality images and results in the laboratory (Merla et al 2010, Tanda 2018), and others producing questionable data in the field (Ganse and Degens 2020). Many other studies provide minimal detail and example images to sufficiently judge the use of IRT during exercise (Buono et al 2007, de Andrade Fernandes et al 2014), de Andrade Fernandes et al (2014), implemented a pause in exercise to image athletes and Hadzic et al (2019) imaged the thigh although during exercise in a static position on the isokinetic dynamometer. This would not be necessary if IRT can accurately measure in motion, highlighting the relevance of the present study. Tanda et al (2018) provided images of limited blurring in the frontal plane, however, they do state that they took 3–4 frames at a time to give multiple image options, indicating movement artefact issues. Furthermore, similarly to Merla et al (2010), they imaged the thigh in the frontal plane on a treadmill where movement across the image is minimised. Images provided by Ganse and Degens (2020) displayed heavy blurring due to movement artefacts and low resolution, producing almost unidentifiable ROIs. This is likely due to their choice of a FLIR ONE Pro Camera, which provides low accuracy, low resolution and longer capture times. Their choice of a low-resolution camera (only 160 * 120 pixels native resolution for the IR sensor). While images are upscaled to 1440 * 1080 pixels when combining it with the visible light camera image, this does not improve the true resolution of the thermal image results in unrealistic skin temperatures, as their ROI selection was a single warmest pixel. Comment on Ganse and Degens (2020) indicate not only the importance of camera selection but the need to analyse the physical constraints of the camera used for research to inform the appropriate application of IRT. The importance of camera type has been further discussed by Machado et al (2021), who found the FLIR One camera to provide poor agreement in skin temperatures, compared to a higher resolution camera (80 * 60 versus 320 * 240 pixels). They however, utilised three low resolution cameras and static participants pre and post exercise.

Herein, the present study provides to our knowledge the first comparison of two high-resolution cameras and specifically during motion. Therefore, when comparing the findings of Ganse and Degens’s (2020) and Machado et al (2021), with our calculations of movement artefact, indicate the importance of implementing a high-resolution camera as the movement artefact is even greater in a low-resolution camera. However, Machado et al (2021) and our data are not compared with a contact method for comparison of accuracy and thus future research should investigate this. In the competition setting of our experiment, this was not possible. Previous literature has compared non-contact and contact methods and found differences in skin temperature, but only once exercise had taken place (de Andrade Fernandes et al 2014, Bach et al 2015b, Maley et al 2020). Post exercise the camera produced temperatures 0.8 °C–1.2 °C lower than thermistors (Priego Quesada et al 2015), however, this could be explained by the skin under the taped sensors changing temperature at a different speed to the nude, exposed skin.

As discussed in detail earlier, blurring is an issue and can be related to various camera characteristics, mainly the time constant (LWIR) and integration time (MWIR) and the image capture method (line scan for LWIR).
versus faster snapshot for MWIR). Running movements are captured well with the fast MWIR, but create problems for the LWIR, with the image moving across pixels during the capture period. This effect is worst when moving laterally through the image but is smaller for areas moving towards or away from the camera in frontal and rear aspect, as observed by Merla et al. (2010) and Tanda (2018). Researchers should identify whether the rate of motion in comparison to the exposure and response times and resolution of the camera compromises the results. This is true of both cyclical (repeated) and acyclical (one off motion) activities. However, the continuous often rapid cyclical movements of running or cycling for example are likely to pose more of a difficulty.

Besides motion, similarity of skin temperature with that of the surroundings as observed in the dataset used, can influence the ability to determine ROIs. Due to the narrow range that human skin temperature can sit in, in general cameras with a wider temperature range, are less sensitive to what are in absolute terms small temperature changes (Zhu and Xin 1999). The distinction between objects can be improved by selecting the appropriate colour palette, such as a high contrast palette. One factor contributing to potential difference in the measurements of each camera, is a constant temperature drift of only the LWIR; caused by a change in the physical temperature of the LWIR detector over time. A regular non-uniformity correction is automatically performed, whereby the shutter between the optic and the detector acts as a reference for the detector to calibrate, during which time frames are lost. Temperatures measured just before and after the shutter calibration may differ slightly due to this drift correction, and this may explain part of the variation in temperatures between cameras.

ROI selection
Obstructions by clothing, other athletes or cooling devices, can prevent collecting a body part or a full area ROI and therefore the importance of the size of the ROI and the impact of missing ROIs required investigation. Other studies do not reference missing data, however, when it seems inevitable in field-based testing the reader should be informed of the magnitude of missing data and how it was dealt with to ensure a representative mean $T_{sk}$ value. Observer bias in the positioning of ROIs has led to debate over how to ensure reliability. The importance of the selection of ROIs should not be underestimated, as it has been shown to influence the reliability of studies (Ring and Ammer 2015). A number of protocols aiming to standardise ROI selection have been proposed, such as the Glamorgan protocol (Ammer 2008), the use of anatomical landmarks or instrumenting participants with markers (Seixas et al. 2020), the use of automated standardised analysis, such as the hottest pixel and the surrounding 24 (Ludwig et al. 2014) or the hottest 5 pixels (Formenti et al. 2017). However, both protocols of Ludwig et al. (2014) and Formenti et al. (2017), create a bias towards analysing the hottest area and lose a high amount of information from the thermogram, negating the benefit of measuring a large skin surface area. This is particularly important where motting is present.

Our results support analysing the largest surface area possible as the results increased in divergence as the number of pixels analysed was reduced, with all the smaller ROI measurements being significantly different from the whole area ROI, except for the line on the LWIR. Although in this instance the average value over participants of a single pixel placed centrally, was within a small error from the average value over participants of the full area ROI, it should be noted from the much higher errors in the maximum and minimum pixels, that the error of a single pixel can be much higher, when left to subjective placement or automated protocols that select the most extreme temperatures. This follows previous suggestions of the poor accuracy of single pixels being augmented in low resolution images (Machado et al. 2021). Analysis relying on the hottest 5 pixels (Ludwig et al. 2014, Formenti et al. 2017) creates a strong positive bias in the representativeness of the skin temperature value, as our results showed the hottest or coolest pixel can deviate by 1.3 ± 0.4 °C from that of the full area ROI. Both cameras displayed similar sensitivities to the size of the ROI, perhaps due to the same NETD at 30 °C and high resolution. The resolution of the camera is very important given that every pixel is a temperature data point; the greater the resolution the greater the amount of information and the fewer pixels need to be discarded at the edges. By positioning a camera closer to the person being imaged, a greater number of pixels within a frame will be that of a person as opposed to surrounding objects. This can help the accuracy of low-resolution cameras, however, as in this case the position of the camera was chosen for logistical reasons and therefore may not provide a solution. Given that ROI selection is one of the most contentious points in IRT any improvement in the measurement of a given ROI is beneficial.

When measuring skin temperature during motion, it would be expected that increased convective cooling from the additional movement of air may result in differing skin temperatures between the anterior and posterior, rendering a contact method or ROI selection that does not cover the three perspectives as less representative of the whole-BSA. Highly variable differences between the three perspectives were recorded in this data set. While the mean differences were not large between perspectives there was large individual variability, and significant differences between the anterior versus lateral and anterior versus posterior, further supporting the greater the surface area analysed, the more representative mean $T_{sk}$. This variation is likely an
effect of the frontal air velocity introduced by moving forward, which is less influential on lateral or posterior skin temperatures; thus, particularly important in field-based or non-static laboratory testing anterior and posterior measures should be taken.

**Sensitivity to missing data**

Although missing data points (ROIs) were inevitable, it was unclear so far whether multiple missing sites would influence the whole body mean $T_{sk}$. The thighs alone had a meaningful impact on the calculated mean $T_{sk}$. Therefore, a person with thigh measurements missing would be excluded from the data set, as a representative mean $T_{sk}$ cannot be obtained. Meaningful determined as 0.5 °C or larger, as used elsewhere (Harper-Smith et al 2010, Bach et al 2015a). All the other body sites, when missing in isolation, did not substantially impact mean $T_{sk}$; however, missing multiple sites did. This supports the need for IRT and other skin temperature measurement to cover the greatest surface area of nude skin possible. Given our observations, well established skin temperatures; thus, particularly important in

**Sensitivity to input parameters**

Input parameters often receive limited acknowledgment in original research and standardisation procedures. This factor becomes even more important in field-testing, where the conditions and distances may not be stable. Analysis software integrates the parameters into the object temperature and therefore the parameters each have an interactive relationship; these relationships and sensitivities are shown in table 3. Cameras displayed greater sensitivity to deviations in emissivity, reflected and air temperatures than relative humidity or distance. One benefit of the LWIR is that due to the high emissivity of skin (0.98) in this wavelength, compared to the MWIR camera (0.91), the sensitivity to deviations in reflected/background and air temperature in this context are much lower. Due to the influence of reflected temperature and air temperature these factors should be measured accurately, particularly in field-based research where they are not necessarily stable.

**Sweat and emissivity**

The appropriate emissivity for the specific camera should be considered by researchers, and not assumed that 0.98, because it is the most used, is the most appropriate. One factor that should be considered in the selection of emissivity is the potential impact of sweat. Typically, it is suggested that skin should be dried prior to imaging (Smith and Havenith 2011, 2012), with a consistent technique and material (Seixas et al 2014). This, however, is not possible during a competitive race, and drying should be advocated with caution due to the lost heat exchange from the sweat that is dried off and therefore not evaporated. Ammer (2009) hypothesised that a layer of sweat on the skin’s surface would act as a filter for infrared radiation and the camera would measure the outer surface of the sweat layer. As discussed in the sweat and emissivity section earlier in detail (penetration depth; changes in emissivity for skin to water in different wavelengths), this is plausible for the LWIR, but not for the MWIR where water is more transparent for IR. In the images collected here, detailed variations of surface temperatures were visible, showing the expected mottling of the skin. Thus, even if the temperature of the thin water layer is taken rather than the actual skin temperature, the camera clearly is still sensitive to regional variation in local skin temperature. Moreover, for heat transfer purposes, it is the temperature of the thin water layer that determines dry and evaporative heat loss, as it determines the surface vapour pressure. For the MWIR it may be assumed, for thin sweat layers, that the value observed in the present data is likely a combination of skin temperature and sweat temperature, while for the LWIR it will be the surface. Given the similar outcomes for both cameras the thin sweat layer during the activity does not seem to have much impact.

**Static versus moving skin temperatures**

As noted, there is limited data on skin temperature using IRT during movement and even more limited data, in an outdoor field setting. In lab sessions, typically movement is briefly interrupted to take a static IR image. In our study, despite the short period of time between finishing the race and the post thermogram, large increases in skin temperature were recorded. A definitive explanation for this cannot be given, however likely contributors include the loss of airflow cooling the skin while moving (lower heat transfer coefficients), removal
of per-cooling, the towelling of sweat from the skin’s surface (reducing evaporation), for the post-thermogram and an end-of-exercise vasomotor response (increasing blood flow due to reduced sympathetic tone). These results bring into question the comparability of skin temperature measured post or between bouts of exercise, for representing in race or exercising skin temperatures and researchers should consider the use of IRT during motion as a more appropriate measurement. One limitation of our data post exercise that should be noted was despite asking participants not to, some were found to douse themselves with cool water prior to towelling dry for the post thermogram. However, if dousing had interfered with the data, it would be expected that skin temperature would have been lower post exercise rather than the observed higher value.

Conclusion

In conclusion, IRT provides a valuable tool for measuring skin temperature of moving athletes; providing field data for those wishing to investigate human thermoregulatory responses, for physiological research, understanding human performance, or heat related safety during athletic competition and/or mass participation events. A cooled quantum detector camera provides greater clarity and reliability of images, when compared to an uncooled microbolometer camera, in a high rate of motion scenario. The present study makes a qualitative and quantitative comparison of these two types of high-resolution camera, for the measurement of skin temperature. The uncooled LWIR camera provides good agreement in skin temperature readings with the cooled MWIR camera, though it suffers substantially more from unusable frames due to movement artefacts, especially in the side aspect. The sensitivity of both cameras, to altering the size of an ROI, missing data and deviations in input parameters were similar. Both cameras showed that the smaller the surface area of the body analysed, the less representative the weighted mean \( T_{sk} \) results are, and therefore ROIs should cover as much BSA as possible, and missing ROIs need to be accounted for. Future research should appreciate the limited comparability of small to larger surface area temperature measurements, and the differences in measurements taken from static participants as a representation of whole-body skin temperature during exercise. In the analysis of the impact of a wet skin surface (sweat), it is evident that the LWIR is hardly affected by emissivity differences between skin and water and that the measured temperature will be that of the surface sweat layer when this developed. The MWIR is affected by a larger emissivity change between skin and water in its wavelength range, but as the water is more transparent in this wavelength (the skin ‘shines through’), the total impact is still limited. The MWIR camera measures a mix of skin and sweat temperature, but with thin sweat layers this distinction may be academic.

As there is not always a consensus on the standardisation of IRT, or possibility to control all the factors that influence the results, it is critical that future literature provides as much methodological information as possible; including type of camera, camera set up, software for analysis, input parameters used, how ROIs were selected, how much data was missing, how this is treated and calculation of mean \( T_{sk} \).

Acknowledgments

The authors thank the medical and scientific department of World Athletic for their endorsement, financial and operational support. The authors thank the local organising committee for their operational support and the following contributors to the data collection of the wider project P-E Adami, J-M Alonso, S Bermon, N Bouscaren, S Buitrago, M Cardinale, C Esh, F Garrandes, J Gomez-Ezeiza, M Ihsan, M Labidi, G Lange, S Moussay, K Mtibaa, L Taylor , N Townsend and M Wilson.

This study was supported by World Athletic (formerly International Association of Athletic Federation, IAAF).

References

Ammer K 2008 The glamorgan protocol for recording and evaluation of thermal images of the human body Thermol. Int. 18 125–9
Ammer K 2009 Does neuromuscular thermography record nothing else but an infrared sympathetic skin response? Thermology International 19 107–108
Bach A J E et al 2015a Does the technique employed for skin temperature assessment alter outcomes? A systematic review Physiol. Meas. 36 pp R27–51
Bach A J E et al 2015b A comparison between conductive and infrared devices for measuring mean skin temperature at rest, during exercise in the heat, and recovery PLoS One 10 e0117907
Buono M J et al 2007 Comparison of infrared versus contact thermometry for measuring skin temperature during exercise in the heat Physiol. Meas. 28 855–9
Buono M J and Ulrich R L 1998 Comparison of mean skin temperature using ‘covered’ versus ‘uncovered’ contact thermistors Physiol. Meas. 19 297–300
Ring E and Ammer K 2015 The technique of infrared imaging in medicine Physiological Measurement 33 R33–R46
Ring E F J and Ammer K 2012 Infrared thermal imaging in medicine Physiol. Meas. 33 R33–46
Robinson P J and Davies J A 1972 Laboratory determinations of water surface emissivity J. Appl. Meteorol. 11 1391–3
Rogalski A 2010 Infrared Detectors 2nd edn (Florida, FL: CRC Press) 9780367577094
Roy R, Boucher J P and Comtois A S 2006 Validity of infrared thermal measurements of segmental paraspinal skin surface temperature J. Manipulative Physiol. Ther. 29 150–3
Sanchez-Marín F J, Calixto-Carrera S and Villaseñor-Mora C 2009 Novel approach to assess the emissivity of the human skin J. Biomed. Opt. 14 024006
Sawka M N, Cheuvront S N and Kenefick R W 2012 High skin temperature and hypohydration impair aerobic performance Exp. Physiol. 97 327–32
Schlader Z et al 2011 Skin temperature as a thermal controller of exercise intensity Eur. J. Appl. Physiol. 111 1631–9
Seixas A et al 2014 A preliminary study on the relationship between energy expenditure and skin temperature in swimming 12th Int. Conf. on Quantitative InfraRed Thermography (Bordeaux, France) pp 90–7
Seixas A et al 2020 Reliability of infrared image analysis based on anatomical landmarks Infrared Phys. Technol. 104 103149
Selie J et al 2006 An accurate and reliable method of thermal data analysis in thermal imaging of the anterior knee for use in cryotherapy research Arch. Phys. Med. Rehabil. 87 1630–5
Shapiro Y et al 1980 Physiological responses of men and women to humid and dry heat J. Appl. Physiol. 49 1–8
Shapiro Y, Pandolf K B and Goldman R F 1982 Predicting sweat loss response to exercise, environment and clothing Eur. J. Appl. Physiol. Occup. Physiol. 48 83–96
Smith C J and Havenith G 2011 Body mapping of sweating patterns in male athletes in mild exercise-induced hyperthermia Eur. J. Appl. Physiol. 111 1391–404
Smith C J and Havenith G 2012 Body mapping of sweating patterns in athletes: a sex comparison Med. Sci. Sports Exercise 44 2350–61
Stekete J 1973 Spectral emissivity of skin and pericardium Phys. Med. Biol. 18 686–94
Tanda G 2013 Skin temperature measurements by infrared thermography during running exercise Experimental Thermal Fluid Science 71 103–113
Tanda G 2018 Total body skin temperature of runners during treadmill exercise: a pilot study J. Therm. Anal. Calorim. 131 1967–77
Togawa T 1989 Non-contact skin emissivity: measurement from reflectance using step change in ambient radiation temperature Clin. Phys. Physiol. Meas. 10 39–48
Tyler C J 2011 The effect of skin thermistor fixation method on weighted mean skin temperature Physiol. Meas. 32 1541–7
van Marken Lichtenbelt W et al 2006 Evaluation of wireless determination of skin temperature using iButtons Physiology and Behaviour 88 489–497
Wingo J E et al 2009 Effect of elevated local temperature on cutaneous vasoconstrictor responsiveness in humans J. Appl. Physiol. 106 571–5
Zhu W P and Xin X R 1999 Study on the distribution pattern of skin temperature in normal Chinese and detection of the depth of early burn wound by infrared thermography Ann. New York Acad. Sci. 888 300–13