2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)

2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) | 978-1-6654-1298-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ISMAR-Adjunct54149.2021.00085

Augmenting Human Perception: Mediation of Extrasensory Signals in Head-Worn Augmented Reality

Austin Erickson * Dirk Reiners † Gerd Bruder ‡ Gregory F. Welch §

University of Central Florida
Orlando, Florida, United States

Abstract

Mediated perception systems are systems in which sensory signals from the user's environment are mediated to the user's sensory channels. This type of system has great potential for *enhancing* the perception of the user via augmenting and/or diminishing incoming sensory signals according to the user's context, preferences, and perceptual capability. They also allow for *extending* the perception of the user to enable them to sense signals typically imperceivable to human senses, such as regions of the electromagnetic spectrum beyond visible light.

In this paper, we present a prototype mediated perception system that maps extrasensory spatial data into visible light displayed within an augmented reality (AR) optical see-through head-mounted display (OST-HMD). Although the system is generalized such that it could support any spatial sensor data with minor modification, we chose to test the system using thermal infrared sensors. This system improves upon previous extended perception augmented reality prototypes in that it is capable of projecting registered egocentric sensor data in real time onto a 3D mesh generated by the OST-HMD that is representative of the user's environment. We present the lessons learned through iterative improvements to the system, as well as a performance analysis of the system and recommendations for future work.

Index Terms:

Computing methodologies—Computer graphics—Graphics systems and interfaces—Mixed / augmented reality; Human-centered computing—Visualization—Visualization systems and tools

1 Introduction

Mediated perception systems are systems that mediate sensory information from the user's environment to the user. While this definition includes technology such as augmented reality (AR) and virtual reality (VR), it also goes beyond to include things such as eyeglasses and hearing aids, which also mediate visual/aural signals in order to enhance the perceptual capabilities of their wearer. Mediated perception systems can be further classified into *enhanced perception systems*, in which the user's range of perception is enhanced in some manner (e.g. glasses, telescopes, hearing aids), and *extended perception systems*, in which the perception of the user is extended beyond the five normal humans senses to include additional sensory channels (e.g. thermal infrared, ultraviolet, abstract spatial data).

Extended perception systems have a wide range of potential applications. For example, access to thermal infrared data has been shown to be useful in the contexts of defense, healthcare, and education, by allowing for easy identification of hidden objects and personnel, identification of fevers and abnormal circulation patterns in patients, and as a method of understanding heat conduction [1, 8, 10].

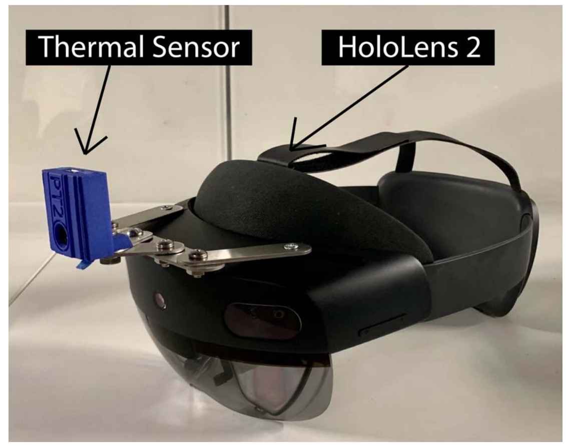

Figure 1: An annotated view of our extended perception system consisting of the HoloLens 2 and an attached FLIR 3.5 radiometric thermal infrared sensor.

AR displays are particularly well-suited for use in mediated perception systems due to their ability to output signals across multiple sensory channels that can be perceived as having originated from within the user's physical environment. While several extended perception systems have been previously investigated that incorporate AR displays into their design (see section 2), these works are often limited in that they are either restricted to video see-through head-mounted displays (VST-HMDs), where the user's visual perception is limited by the resolution of the display, or they incorporate sensors placed in fixed positions within the environment, which prevents portability and explorative use cases. These approaches also offer few ways to explore the extrasensory data, such as through filtering of extrasensory data by user-defined ranges of interest, or through user-defined visualization modes.

In this paper, we present a prototype extended perception system that addresses these limitations. The system allows for data from a USB sensor to be streamed to an optical see-through head-mounted display (OST-HMD), the Microsoft HoloLens 2, where it can be filtered according to the user's preferences, projected onto the spatial mesh created by the HMD, and presented to the user in stereoscopic 3D under several different visualization modes. While this system can be easily modified to accept any spatial USB sensor, we chose to test it with an FLIR 3.5 radiometric thermal infrared sensor, which allows the user of the system to view the thermal properties of their surroundings. We present our approach to implementing this prototype system, a performance analysis of the system, and provide recommendations for resolving the system's current limitations.

The rest of the paper is structured as follows: section 2 provides additional background and relevant work at the intersection of medi-

*e-mail: ericksona@knights.ucf.edu
†e-mail: dirk@ucf.edu
‡e-mail: bruder@ucf.edu
§e-mail: welch@ucf.edu

978-1-6654-1298-8/21/$31.00 ©2021 IEEE
DOI 10.1109/ISMAR-Adjunct54149.2021.00085

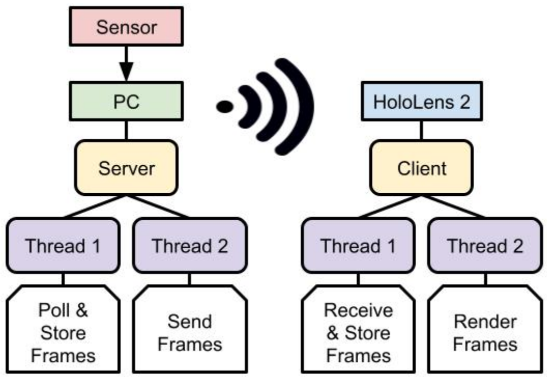

Figure 2: Block diagram illustrating our extended perception system.

ated perception systems and AR. Section 3 describes our prototype system. Section 4 describes a performance analysis run using the system, and section 5 concludes the paper.

2 Related Work

Several mediated perception systems have been implemented and investigated in prior research. For example, Orlosky et al. created a prototype system called VisMerge, in which thermal infrared data is mapped into visible light, combined with visual data from the user's environment, and is presented to the user via a VST-HMD [11]. Employing VST-HMDs in mediated perception systems come with several pros and cons. With these displays, virtual imagery is less prone to change in response to environment conditions, as for example optical see-through displays can have difficulty achieving adequate contrast in brightly lit environments [4]. However, the user experiences a lower resolution view of their physical environment and a more restricted field of view compared to their natural vision without the HMD, since all visual information is being mediated to them through a near eye display with a limited resolution and size. While display technology may eventually advance to overcome this limitation, a more immediate solution is to use either OST-HMDs or spatial AR when investigating mediated perception systems.

The use of spatial AR for mediated perception systems was investigated previously by Cao et al. [3]. They built a system in which thermal data is gathered from several fixed-position sensors in an environment and is then presented in the environment via use of projectors. This approach allows the user to not have to wear an HMD, however due to the sensors being placed in fixed positions, it is more suited for specific applications in which the thermal flux of an object or scene needs to be closely monitored over time, and it is less suited for exploration using the extended perception system.

Knierim et al. investigated thermal visualization using an OST-HMD, the Microsoft HoloLens 1 [8]. Their prototype allows the user wearing the HoloLens to visualize thermal data from a nearby sensor that is also set in a fixed position in order to observe the thermal properties of a heated rod. The system uses vuforia markers to align the visualization in 3D space over the position from which the data is originating.

These two previous approaches are both somewhat limited in that is difficult for a user to be mobile while taking advantage of their extended range of perception. While this may not be an issue for certain tasks, should the user need to explore an environment while using the extended perception system (e.g. performing a thermal efficiency inspection of a building), then an alternative approach is needed.

Erickson et al. previously created a prototype mediated perception system that uses an OST-HMD, the Microsoft HoloLens 1 [6]. Their system used two thermal sensors mounted to the HoloLens that provide a stereoscopic 3D thermal vision overlay of the user's environment. This prototype was later extended to additionally incorporate a pair of ultraviolet sensors, resulting in a system in which the user could toggle between no visual overlay, a thermal vision overlay, and an ultraviolet vision overlay [5]. This prototype also incorporates a backpack-style computer, so that the sensors, HMD, and computer are all fixed to the user, and they are able to freely walk around their environment while using the extended perception system. The main limitations with their approach are that the imagery from the sensors is presented in stereoscopic 3D such that the entire field of view of the display is used, and the images are not registered to the user's environment. This approach also involves pairs of identical sensors to generate the stereoscopic imagery observed by the user, which can be expensive and may introduce perceptual drawbacks depending on the discrepancy between the users inter-pupillary distance and the distance between the pairs of sensors [2]. In the previous work by Erickson et al. these sensor pairs are placed at distances significantly greater than a typical inter-pupillary distance of the user, which creates an exaggerated sense of depth for imagery within the binocular overlap of the system. Additionally the system creates a large region within the user's personal space in which the imagery does not fall into the binocular overlap region of the sensors, making it difficult to observe sensor imagery near to the user.

The prototype described in the next section extends the work by Erickson et al. by reducing the number of sensors needed by projecting the sensor data onto the 3D reconstruction of the user's environment that is generated by the OST-HMD. Since a single sensor is used, the limitations secondary to binocular overlap and the discrepancies between the sensor placements and the user's inter-pupillary distance are overcome. In this manner, the user is able to observe real-time thermal imagery that is registered to the user's environment.

3 Prototype System

In this section, we describe our prototype extended perception system that builds off of the previous work described in section 2.

When determining which display to use in our system, we were motivated by a brief analysis of the current use cases of spatial extrasensory information. We noticed that most applications involve either locating sensor signals that stand out from the rest of the scene, such as in defense and security applications, or identifying change in sensor signals over time, such as in industrial monitoring applications. This means that there are certain sensor ranges for these applications that are not of interest to the user, and that could be ignored in order for the user to retain a less restricted and natural resolution view of their physical environment.

These applications are also very context specific, so they would likely account for a relatively small percentage of use cases for an extended perception system, and would otherwise not be a large factor contributing towards the widespread adoption of extended perception systems by society. Because of this, the type of display chosen for the extended perception system should be based on the ubiquitous devices that users will have constant and convenient access to in the future. While at the current time, these devices would be smart phones, we believe that OST-HMDs may eventually take their place due to their ability to present virtual information directly into the user's environment rather than onto a small handheld display. Because of this, we opted to use the newest commercial OST-HMD available to us, the Microsoft HoloLens 2.

As previously mentioned, our extended perception system is designed so that it could be used with any off the shelf USB spatial sensor. We chose to use a FLIR 3.5 radiometric thermal infrared sensor, housed in a PureThermal 2 breakout board for testing with our mediated perception system. This sensor has a spectral range

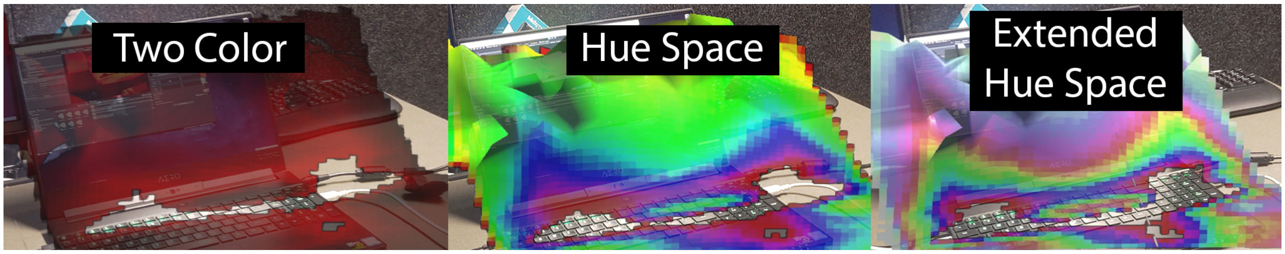

Figure 3: This image was taken via screen capture on the HoloLens 2 and depicts the three visualization modes available for use in the prototype extended perception system across a temperature range of 27 to 38 degrees Celsius.

between 8 and 14 micrometers, a resolution of 160 by 120 pixels, and a diagonal field of view of 71 degrees. The sensor was mounted to be centered above the integrated RGB camera on the front of the HoloLens 2 via the use of several interconnected brackets (see figure 1).

As there is currently no way to directly connect off the shelf sensors to the HoloLens 2, the block diagram of our system (shown in figure 2) consists of several nodes. The sensor is connected via USB to a host PC, which runs a server application that is responsible for gathering data from the sensors and using a UDP network connection to stream the data over a local network. The server is written in C#, and uses the Lepton windows SDK to handle the initial setup of the thermal sensor, where radiometric mode is enabled and a callback function for receiving incoming frames is set up in a new thread. In radiometric mode, incoming data is received from the sensor is in the form of a 160 by 120 array of 16-bit unsigned short temperature values in units of Kelvin multiplied by 100. This raw temperature data is passed as a byte array into a queue that is shared by the main thread. In the main thread, the latest available sensor data is read from the queue and sent across the network to the client, then all other entries are removed from the queue.

The client was developed with Unity and C#, and runs on the HoloLens 2 (running the windows holographic OS version 21H1, build 20346.1002). The unity application deployed (or streamed via holographic remoting) to the device accepts incoming data from the server and stores it in a queue in one thread, while another thread is continuously passing the latest data in the queue as an R16 texture into the appropriate custom shader for the current rendering mode. The shader first restores the value to its original bits by scaling the value by $2^{15} - 1$, after which it performs a conversion of units into Celsius or Fahrenheit, and examines a minimum and maximum temperature range for incoming data which are statically defined in the shader properties. Pixel values within this range are rendered according to one of several rendering modes (described below), while values outside the range are rendered as transparent.

Unity version 2019.4.26 was used to implement everything on the client/HoloLens 2 side. Within the Unity environment, a gameobject is parented to the gameobject that represents the position/orientation tracked HMD. This child object is used to represent the position of the sensor relative to the HMD within the scene, and this position was manually entered from direct measurements made on the device. On this gameobject, a unity projector component is added and is configured according the field of view and aspect ratio parameters of the attached sensor. A Unity material is also set up to use the shader for the desired rendering mode (described below) and is attached to this projector component so that the texture passed into the shader is projected out into the unity scene. In this manner, the projection is registered to the user's environment in real time according to the current position and orientation of the HMD.

The Microsoft mixed reality toolkit is also loaded into the Unity scene and is configured so that Unity receives the spatial mapping data that is generated and updated in real time by the HoloLens 2. The parameters are set so the mesh is updated every 1.75 seconds, and is set within bounds of a three meter cube. This current update rate was reduced from its default value 3.5 seconds with the intention that dynamic objects in the scene should be more easily captured and included into the spatial mapping data.

There are three different visualization modes available to the user that are implemented via custom shaders. In the *Two Color* mode, a simple linear interpolation is performed in RGB color space between two user-defined colors. The left-most image in figure 3 shows an interpolation between black/transparent cool temperatures and red warm temperatures. This mode allows for easy identification of parts of the scene that fall within the user defined sensor ranges, but does not offer a large range of colors for investigating details such as temperature gradients on the included objects. In the *Hue Space* mode, the interpolation is performed within hue/saturation/value (HSV) color space across all hues while keeping saturation and value fixed at one (maximum) (figure 3 middle). This rendering mode has an increased range of colors that allows the user to visualize sensor gradients within the user-defined sensor ranges, however it is not as intuitive as the two color mode for identifying specific sensor values within the scene. In the *Extended Hue Space* mode, the interpolation is again performed in HSV color space, but instead spans across all hues in the HSV color space twice while simultaneously traversing across all values in the HSV color space once (figure 3 right). In other words, minimum sensor values within the user-defined range are set to HSV triple (0,1,0), whereas sensor values at the midpoint of the range are set to HSV triple (1,1,0.5). Just after this midpoint, the triple resets the hue component such that the HSV triple is (0,1,0.5), and then finally sensor values at the maximum range are set to HSV triple (1,1,1). This mode offers an increased color space over the previous-described modes, which enables visualization of sensor gradients in even more detail, however estimation of sensor values within this space is even less intuitive.

4 Performance Analysis

In this section, we present an analysis of the extended perception system's performance in terms of latency and frame rate.

4.1 Analysis of Frame Rate

Frame rate was measured using the mixed reality toolkit's built in diagnostic system, which overlays frame rate and memory usage data into the user's field of view on the HoloLens 2. The main application was built in release mode in Microsoft Visual Studio 2019 and was deployed to the device, where the analysis revealed that the system meets Microsoft's recommended frame rates for HoloLens applications of 60 frames per second [1].

[1] https://docs.microsoft.com/en-us/windows/mixed-reality/develop/platform-capabilities-and-apis/understanding-performance-for-mixed-reality

4.2 Analysis of Latency

We measured the latency of the sensor imagery employed in our system by analyzing a video recording made via the screen capture software built in to the HoloLens 2. We recorded a video in which the user looks ahead while keeping their head still, and then makes short repeated movements with stationary gaps in between. Latency was measured in video editing software by measuring the number of frames that elapsed between the user moving their head and the sensor imagery updating in response to the movement. From this, it was found that the system suffers from significant end-to-end latency issues on updates to the sensor imagery, which was measured to be 530 milliseconds.

This latency value is near to what was experienced by Orlosky et al. when they tested their VisMerge extended perception system without the use of their custom timewarping algorithm to synchronize the thermal infrared imagery to the RGB camera imagery based on the latest position and orientation of the HMD [11]. In their work, they are able to effectively eliminate this latency, reducing it to 2.6 milliseconds, by using their timewarping approach. Latency perceptual thresholds for AR HMDs were examined previously by Jerald and Whitton, where they made recommendations that latency should be under five milliseconds in order to achieve imagery that has no perceptible latency [7]. From this, it seems likely that if a time-warping approach similar to the one employed by Orlosky et al. is applied to the next iteration of this prototype, then the latency in the updates to sensor imagery could be reduced to acceptable levels.

5 Discussion

In its current state, there are several limitations to our prototype extended perception system. In this section, we will discuss them and present some options for addressing them in future work.

5.1 Spatial Mapping

Our system uses the built-in spatial mapping capabilities of the HoloLens 2 to create a 3D reconstruction of the user's environment upon which sensor imagery can be projected. In its current form, the mesh generated by the device has good correspondence with the physical environment for walls and floors, however the system has difficulty generating a mesh that corresponds well to small or irregular shaped objects. This is further complicated by the way the system handles its updates to the mesh, where an object must be static for multiple update cycles until it begins to be integrated into the spatial mesh. This makes it difficult to project sensor imagery onto small or irregular objects as well as difficult to observe dynamic people and objects with the extended perception system.

In order to solve these issues, it is likely that researchers designing future extended perception prototypes may have to consider using alternative approaches to the spatial mesh, especially if the system needs to be capable of projecting sensor data to dynamic scene elements. One potential alternative would be to make use of direct readings from the RGB-D data from the on board sensors of the HoloLens 2. With this approach, sensor imagery could be more easily mapped onto dynamic objects and it would offer a straightforward method of storing sensor data relative to its source position within the 3D environment. However, this approach presents new challenges, as it would require a precise correspondence between the sensor of interest and the depth sensors on the HoloLens, and an approach similar to Orlosky's timewarping algorithm would be needed to minimize latency between the two sensors [11].

5.2 Registration and Calibration

As mentioned in section 3, the registration between the sensor imagery and the user's view of their physical surroundings was handled by taking measurements of the sensor's position relative to the HMD, and positioning the Unity projector object to the same relative position and orientation. While this approach is useful for quick prototyping and debugging, the registration achieved by this method is less accurate than an alternative approach involving computer vision, and this is readily apparent when observing objects at distances beyond several meters of the user, which may be problematic depending on the user's context. Because of this, consumer-ready extended perception systems will need to ensure that registration error is minimized by employing more robust techniques to calculate the position of the sensor in real time [9]. This however, is a difficult problem to overcome for some sensors that may be included in future extended perception systems, as for example thermal sensors cannot see a standard checkerboard pattern used to calibrate a normal RGB camera, so alternative calibration processes must be used [12].

5.3 Storing Sensor Data

In its current state, our prototype system allows for the user to visualize sensor data in real time, however it does not offer a method of pairing this data with its position of origin within the user's physical environment. By pairing the data with its position of origin, many options are opened up that allow the system to analyze sensor data over time, and offload the identification of spatial or temporal sensor readings of interest from the user to the system. In this manner, the user may receive notifications if a new sensor signal of interest appears in the scene or if an existing sensor signal has changed over time to be significantly different than usual. We imagine this being useful not only for professional applications such as security, defense, and healthcare, but also for helping users avoid dangerous situations and health concerns in their daily life [1, 5, 10].

6 Conclusion

In this paper, we have presented a prototype extended perception system that allows the user to visualize, filter, and explore environment-registered thermal sensor data in real time using the Microsoft HoloLens 2. This prototype goes beyond the existing work in the field in that it is capable of projecting sensor data from a single sensor onto the spatial mesh representing the user's physical environment, and the data is capable of being filtered and rendered according to the user's preferences and needs. We have demonstrated the steps taken to realize this current implementation, we have discussed the system's capabilities and limitations, and we have provided recommendations on how to improve the design for future extended perception systems.

Acknowledgments

This material includes work supported in part by the National Science Foundation under Award Number 1564065 (Dr. Ephraim P. Glinert, IIS) and Collaborative Award Numbers 1800961, 1800947, and 1800922 (Dr. Ephraim P. Glinert, IIS) to the University of Central Florida, University of Florida, and Stanford University respectively; the Office of Naval Research under Award Numbers N00014-17-1-2927 and N00014-21-1-2578 (Dr. Peter Squire, Code 34); and the AdventHealth Endowed Chair in HealthcareSimulation (Prof. Welch).

References

[1] A. Akula, R. Ghosh, and H. K. Sardana. Thermal imaging and its application in defence systems. vol. 1391, pp. 333–335, 2011.

[2] G. Bruder and F. Steinicke. Perceptual evaluation of interpupillary distances in head-mounted display environments. In *Proceedings of the GI-Workshop VR/AR*, pp. 135–146, 2011.

[3] Y. Cao, Y. Dong, F. Wang, J. Yang, Y. Cao, and X. Li. Multi-sensor spatial augmented reality for visualizing the invisible thermal information of 3d objects. *Optics and Lasers in Engineering*, 145:106634, 2021.

[4] A. Erickson, K. Kim, G. Bruder, and G. F. Welch. Exploring the limitations of environment lighting on optical see-through head-mounted displays. In *Symposium on Spatial User Interaction*, SUI '20. Association for Computing Machinery, New York, NY, USA, 2020.

[5] A. Erickson, K. Kim, G. Bruder, and G. F. Welch. Beyond Visible Light: User and Societal Impacts of Egocentric Multispectral Vision. In J. Y. C. Chen and G. Fragomeni, eds., *Virtual, Augmented and Mixed Reality*, pp. 1–19. Springer Nature, 2021.

[6] A. Erickson, K. Kim, R. Schubert, G. Bruder, and G. Welch. Is it cold in here or is it just me? analysis of augmented reality temperature visualization for computer-mediated thermoception. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 202–211, 2019.

[7] J. Jerald and M. Whitton. Relating scene-motion thresholds to latency thresholds for head-mounted displays. In *2009 IEEE Virtual Reality Conference*, pp. 211–218, 2009.

[8] P. Knierim, F. Kiss, and A. Schmidt. Look inside: Understanding thermal flux through augmented reality. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 170–171, 2018.

[9] E. Marchand, H. Uchiyama, and F. Spindler. Pose estimation for augmented reality: A hands-on survey. *IEEE Transactions on Visualization and Computer Graphics*, 22(12):2633–2651, 2016. doi: 10.1109/TVCG.2015.2513408

[10] J. Mercer and E. Ring. Fever screening and infrared thermal imaging: Concerns and guidelines. *Thermology*, 19:67–69, July 2009.

[11] J. Orlosky, P. Kim, K. Kiyokawa, T. Mashita, P. Ratsamee, Y. Uranishi, and H. Takemura. Vismerge: Light adaptive vision augmentation via spectral and temporal fusion of non-visible light. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 22–31, 2017.

[12] S. Prakash, P. Y. Lee, T. Caelli, and T. Raupach. Robust thermal camera calibration and 3D mapping of object surface temperatures. In J. J. Miles, G. R. Peacock, and K. M. Knettel, eds., *Thermosense XXVIII*, vol. 6205, pp. 182 – 189. International Society for Optics and Photonics, SPIE, 2006.