Texture Estimation Using Thermography and Machine Learning

Tamas Aujeszky
Engineering Department
New York University Abu Dhabi
Abu Dhabi, United Arab Emirates
tamas.aujeszky@nyu.edu

Georgios Korres
Engineering Department
New York University Abu Dhabi
Abu Dhabi, United Arab Emirates
george.korres@nyu.edu

Mohamad Eid
Engineering Department
New York University Abu Dhabi
Abu Dhabi, United Arab Emirates
mohamad.eid@nyu.edu

Farshad Khorrami
Electrical and Computer Engineering Department
New York University
Brooklyn, United States of America
khorrami@nyu.edu

***Abstract*—Contactless material characterization has the potential to be used in various applications such as teleoperation and autonomous physical interaction robotics. Active infrared thermography is a promising approach for classifying materials based on their thermal response to laser excitation over a short distance, thus creating a contactless haptic modeling scheme. However, factors such as the texture of the object under inspection can influence the thermal signature and therefore need to be compensated against. This paper presents a method to use the exact components of a thermographic material characterization system to estimate texture, allowing it to produce more robust characterization in the presence of textured surface. Experimental results confirm that the system is capable of estimating the texture of the sampled material surface to a sufficient degree, with a promising outlook for further improvements as the data set is scaled.**

***Index Terms*—Texture Estimation, Neural Networks, Active Thermography**

I. Introduction

Advancements in computer vision technologies, neural networks and processing power have together enabled the development of autonomous robots and vehicles that possess decent perception of their immediate unknown environments. However, these technologies are still far from achieving their full potential. For instance, an autonomous robot faces a problem when interacting with unknown objects when it comes to physical interaction due to the challenge of calculating the contact or gripping force. If the amount of force exerted on the object is too much, it could permanently damage the object. On the other hand, applying a gripping force that is too little might result in the object slipping out of the robot's end effector. Therefore it is important for the robot to have some knowledge of the unknown object's physical properties before it comes into contact with it.

With virtual reality the current state of the art is a high definition playback capability of the audiovisual information from real-life scenes. Experiencing 360 degree video recording significantly enhances visual fidelity. However, the sense of immersion is hindered by the lack of ability to physically interact with any of the objects in recorded scenes. Haptic interfaces that could provide appropriate interaction forces are available, however modeling/recording the haptic properties of objects in these scenes remain a challenging task.

Active infrared thermography is a promising technology to allow for non-contact measurements of material physical properties, but by itself it is not robust to factors such as texture and emissivity. This paper presents a method to use infrared thermography to estimate the texture of an object in a contactless fashion. This is achieved by subjecting the object to a short thermal excitation and recording its cool-down with a thermal camera. The data is fed into a multi-channel neural network which performs regression to a numerical value describing the texture depth of the object. Due to the equipment used, this process has the potential to form a vital part of a thermography-based contactless material characterization system (without any need for additional hardware), by feeding the estimate of the texture of the object as an auxiliary input to a neural network. This allows the system to compensate for this factor and training it on objects with a large variety of textures creates a robust system that is capable of working with objects that have a realistic range of textures.

The rest of this paper is organized as follows: section 2 introduces the background of contactless material characterization, why texture estimation plays an important role in the development of a thermography-based haptic mapping sensor, and how this proposed method fits into the Haptic Eye framework. section 3 describes our experimental methodology for acquiring the thermal data and training a custom-designed neural network to produce an estimate for the texture. The results are shown in section 4 and a discussion of them is laid out in section 5. Finally, the conclusions and our future research aims are stated in section 6.

978-1-5386-8344-6/19/$31.00 ©2019 IEEE

II. Background

The field of infrared thermography is oriented around the thermal radiation that all objects emit. The most frequent use of thermography is qualitatively examining thermal gradients on various surfaces [1], [2]. When this gradient occurs without interfering with the system under observation (such as in a surveillance footage or with thermal leakage), it is referred to as passive thermography. On the other hand, active thermography examines the thermal signatures of induced thermal gradient. The means to induce this thermal gradient can be optical (by shining light on the surface), mechanical (by vibrating the surface) or electromagnetic (using eddy currents). The duration of the excitation can be infinitesimal (flash thermography), continuous constant (step thermography) or continuous periodical (lock-in thermography). An overview of these variations is presented in [3].

Ever since its inception in the 1980s, thermography has been mainly used for nondestructive testing and evaluation [4]–[13]. However, some use cases in examining works of art [14]–[17] and testing of electronic circuit boards and components [18]–[20] are also notable. The process of heat flow inside a material is explained by the heat equation [21]. Therefore examining the spatial/temporal evolution of a thermal gradient gives information about the material itself. This idea has been used to estimate thermal diffusivity in a number of studies [22]–[28] using a closed form solution to the heat equation that applies for instantaneous Gaussian excitation.

In our previous work, we presented a framework for characterizing unknown material without making any physical contact between the sensing device and the sampled object [29]. The proposed framework is validated for contactless material characterization using machine learning algorithms in [30] to perform classification between different objects based on their thermal signature. A diagram of this approach can be seen in Figure 1.

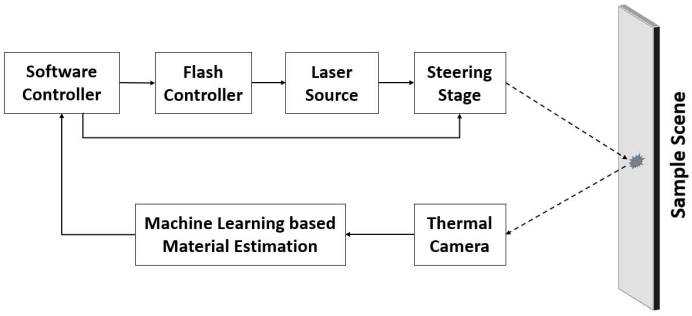

Fig. 1. Diagram of the contactless material characterization approach.

Under this framework, the a Software Controller element is responsible for supplying the desired excitation waveform information to the Flash Controller element and controlling the Steering Stage to enable scanning multiple objects or potentially multiple points on the surface of an object. The Flash Controller instructs the Laser Source to perform the excitation according to the received parameters (duration, power, frequency, etc.). A suitable part of this excitation process is recorded by the Thermal Camera component and is sent to the Machine Learning element which performs processing on the received data and then feeds it to a pre-trained machine learning classifier or regressor. This Machine Learning element can scale from multi-class SVM or KNN learners to complex neural networks. Depending on the applicational needs, it can perform classification between different classes of materials or regression towards various physical properties (such as thermal diffusivity or density) of the object.

In order to create a thermography-based system that is robust in all realistic conditions a key factor must be compensated for: the texture of the material. The texture of the sample acts as a complex factor as it locally distorts the flow of the heat based on the curvature of the sample surface at each point. The objective of this study is to investigate the effects of texture on the characterization of material properties, and to estimate the texture depth in order to make material characterization texture-independent.

III. Experimental Methodology

A. Experimental Setup

The experimental setup is visible in Figure 2. It consists of the following components: a Xenics Gobi-640-GigE thermal camera, a 405 nm laser diode (US-Lasers Inc., model D405-120) and an Arduino board to operate a relay for the laser diode. The thermal camera is set to operate in radiometric mode, providing a 16-bit image with VGA resolution (640x480) at a frame rate of 50 Hz. This setup is completed by a desktop computer that is connected to the camera through a Gigabit Ethernet connection (both instructing the camera and acquiring the recorded frames) and to the Arduino board via a serial connection. The data acquisition process is controlled from a program in the MATLAB R2018a environment. The sample is 195 mm away from the camera and 225 mm away from the laser diode.

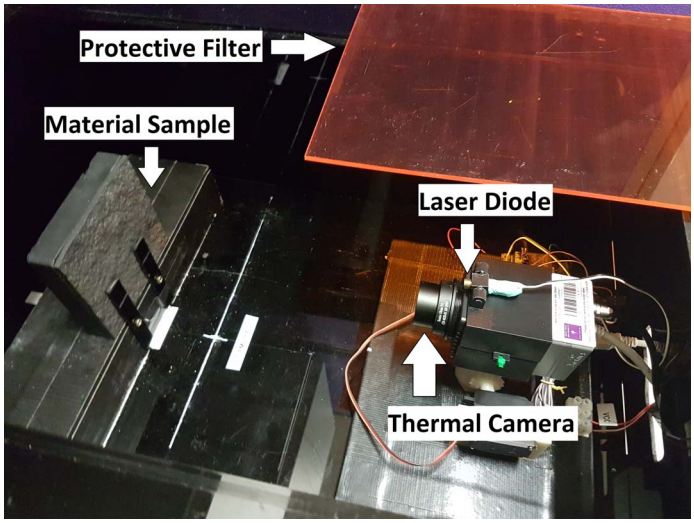

Fig. 2. Experimental setup.

A total of 15 experimental samples were used, with varying textures. These samples were all 3D-printed from the same material, using a MATLAB algorithm for producing random fractals with Hurst exponent values of 0.2 (fractal dimension of 2.8) and standard deviations of 0.1 mm to 1.5 mm in texture depth, corresponding with the numbering of the samples. This covers a large range in which most of the everyday objects and materials fit into. Figure 3 depicts our sample set with #1, and #15 acting as the smoothest and the roughest endpoints of our texture scale, correspondingly.

B. Data Acquisition

Data acquisition for a single experimental round consists of the following steps: first, the thermal camera records 40 radiometric frames to be used as the ambient baseline (these are called 'ambient' frames). Next, the laser diode is turned on for 5 seconds, during which it heats the sample to a slight extent. The laser diode is then turned off and the camera immediately captures the first 200 frames (approximately 4 seconds) of the cooldown process (these are the 'cooldown' frames). The recording is processed by taking the per-pixel average of all ambient frames and subtracting them from each cooldown frame. Finally the 41x41 pixel environment of the excitation center (the highest value pixel) is cropped for each processed frame and these frames make up the data set.

Once these steps are over, a two-minute timeout is applied to ensure that the sample returns to thermal equilibrium before starting again. The set of these steps comprises one experimental round. During the data acquisition process, 5 experimental rounds were carried out on each sample before exchanging them to the next one. Performing these 5 rounds on all 15 samples made up a single experimental session. In order to collect enough data, a total of 20 sessions were carried out, resulting in a total of 300,000 processed frames. The placement of the samples was slightly altered in the lateral direction between different sessions to ensure that the excitation hits different points of the surface samples in different sessions.

C. Neural Network Design and Training

Given that our input data to the network is a set of frames, the use of a convolutional neural network (CNN) architecture was the most fitting starting point in our design process. When designing the neural network for performing regression towards texture level we needed to consider ways to overcome the challenge of bridging the gap between the large number of trainable parameters of neural networks and the size of our data set. Modern convolutional neural networks have upwards of millions of parameters [31] and require even more data entries for training in order to avoid an underdetermined network that is not guaranteed to generalize well.

The first decision to bridge the gap between the number of our recordings and the parameter number of the network is to use single frames as separate data points instead of using an entire recording as a single entry (while ensuring that different frames from the same recording can only be present in the training or the validation partition, but not both at the same time). On one hand this ensures we have a data set size of 300,000 instead of 1,500. On the other hand, using single frames to train the network makes it a much more difficult task. In particular, the network will suffer to categorize two separate frames taken at the beginning (frame #1) and the end (frame #200) of the same recording as belonging to the same sample. In order to overcome this problem we decided to modify the standard CNN architecture and insert the time stamp of the frame (defined as the difference in time between the end of the excitation and the capturing of the frame) as an auxiliary input to the network right after the convolutional stage and before the fully connected stage. This ensures the network has potentially enough information to compensate for the difference between earlier and later frames. A diagram of this architecture can be seen in figure 4, with the time stamp auxiliary input at the right side parallel to the convolutional stage. This network has a total number of trainable parameters of 5577, and is therefore suitable for the size of our data set.

An additional problem for our texture recognition method is that the captured frames can vary greatly based on whether the center of excitation falls at a point on the surface of the object that is flat (zero curvature), bulging out (positive curvature) or cratering in (negative curvature). The curvature of the surface directly affects the heat flow and it poses the biggest problems for rougher surfaces, where points with both low negative and high positive curvature values are present. One reasonable assumption is that sampling the same sample at a set number of different points on its surface (with the previously defined excitation pattern) and taking the acquired recordings together can provide a more robust view of the surface than using a singe footage. To this end, we generalized the previous neural network with a variable number of input branches, corresponding to the number of recordings performed. Each of these recordings goes through a separate (but structurally identical) convolutional stage. Each branch network ends in a single dense layer and these are fed into another dense layer before the output stage.

Given our lack of enough equipment to do simultaneous excitation at multiple locations on the surface of the sample, this was overcome by feeding the frames of a number of different recordings of the same sample into the network at a time, through separate convolutional "branches", all concatenated right before the last fully connected layer of the network. While doing this process it was ensured that the network is only fed frames with the same order number (e.g. frame #58 for each input) at the same time during both training and validation. Given that different sessions corresponded to different excitation locations on the sample surfaces, this technique is suitable to replicate simultaneous sampling at different points on the surface. We have built and tested networks with 1, 2, 4 and 8 branches separately to gain insight on the usability of this technique. Finally, we have also used data augmentation to increase the data size and evaluated its effect on the training results. We have used 2x and 4x augmentation levels where transformations included

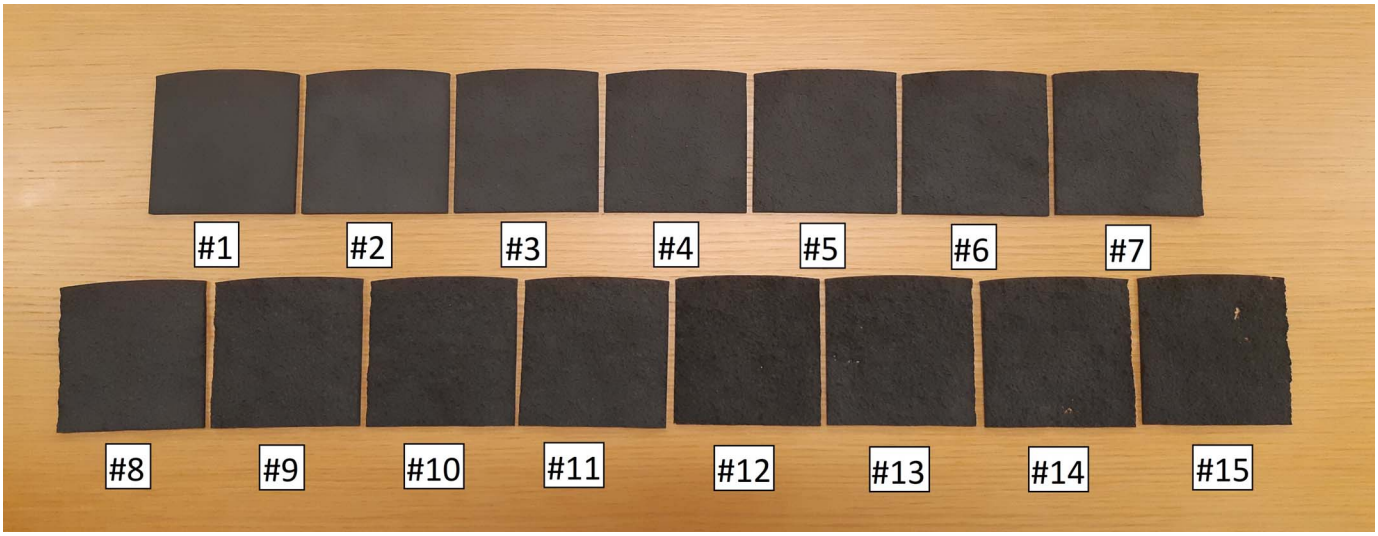

Fig. 3. Experimental samples, ordered from #1 (smoothest, top left) to and #15 (roughest, bottom right).

180 degree rotations and 90 degree rotations, respectively.

The network was trained using 10-fold cross validation by organizing the recordings based on their session number to 10 random partitions with 2 sessions in each. For each fold, 9 of these partitions (18 sessions) were used for training the network and the remaining partition was used for validating the performance of the network. The Adam optimizer was selected with a learning rate of 0.01 for training the network, with mean square error as the loss function. The learning rate was reduced by a factor of 2.5 after every 6th epoch that did not yield an improvement, and it concluded after 25 consecutive epochs with no improvement in the validation loss. The regression target values were 1-15 corresponding to the serial number of the samples, which is also corresponding to the standard deviation in the height differences over the surface of the samples (0.1 mm - 1.5 mm, correspondingly).

IV. Results

We trained networks consisting of 1, 2, 4 and 8 branches of inputs with data augmentation levels of 1x (no augmentation), 2x and 4x. We chose to evaluate the results based on the R^2 value of the target values in the validation data set against the values that were predicted by the trained network based on their corresponding recordings. This value captures how much of the variance in the data set is "explained" by the model - a perfect fit achieves 100% score, while a score of 0% is as good as guessing the mean of the target data set (in this case, 8) for each sample. Table I presents the results as the average for all folds in the cross-validation process.

There are two trends that can be deduced from this table. First, increasing the factor of augmentation increases the validation R^2 value of the network. Second, increasing the amount of branches also increases the validation R^2 value, to a certain point. For 1x augmentation, 4 branches yields the highest validation R^2 value and 8 branches performs notably worse. For 2x augmentation, 4 branches is still the best configuration option but the 8 branches option performs nearly as well. However for 4x augmentation, the 8 branches overtakes the 4 branches option as the optimal selection with the highest validation R^2 value, and becomes the overall best configuration option in this table with an R^2 value of 76.629%.

TABLE I
Validation R^2 values

	1 branch	2 branches	4 branches	8 branches
1x augmentation	60.956%	61.184%	64.907%	59.912%
2x augmentation	62.366%	69.778%	70.977%	68.442%
4x augmentation	64.846%	74.887%	75.519%	76.629%

R^2 values for the validation data set, averaged over 10 folds, using networks with 1, 2, 4 and 8 branches and augmentation levels of 1x, 2x and 4x.

V. Discussion

These results clearly demonstrate that both the proposed process and the proposed network architecture are viable for texture estimation. The trends for increasing the level of augmentation and the number of branches show that increasing the complexity of the network and feeding it with a larger data set both improve the R^2 values. Given that the parameter number for this network even with 8 branches is still comparatively low with respect to currently used neural networks for general image recognition, and that our data set can also be scaled further, this implies that these results can undergo additional improvements.

It is interesting to discuss the eventual data path of the proposed thermographic framework when the above presented texture compensation approach is applied to it. This would involve the texture value as output of the multi-branch texture estimation neural network to be fed into an auxiliary input of

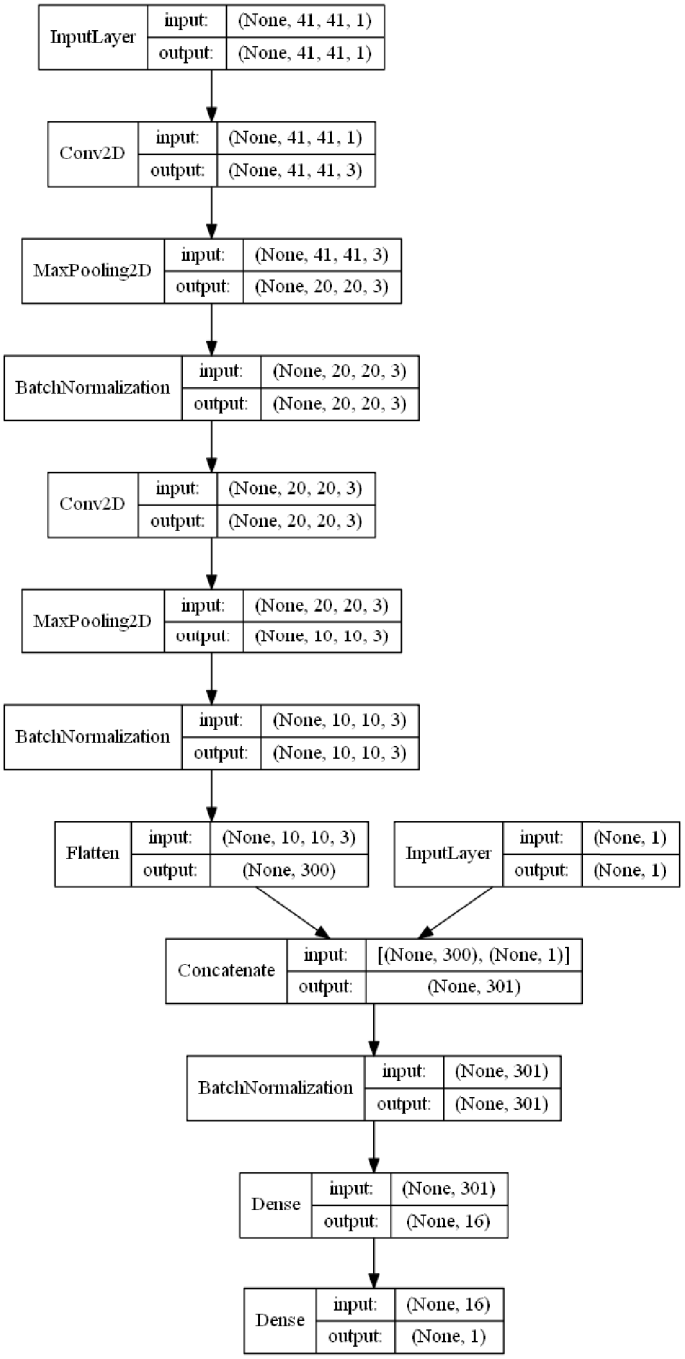

Fig. 4. Single-branch modified CNN architecture with auxiliary time stamp input. Multi-branch variants of this network have multiple parallel convolutional stages that are concatenated before the final dense layer.

the material characterization network (based on an identical structure), parallel to the time stamp information. This means that the overall network would have to be trained in two steps. First by training the texture estimation network and then training the overall network with the former as a fixed weight subsystem. However, despite this and the asynchronous data flow during inference, it is still possible to conduct only a single scanning process, since both the texture subnetwork and the characterization network use the same input data.

VI. Conclusion and Future Work

This article presents a way to estimate texture using the components of a thermographic material characterization system to enhance its immunity against varying levels of texture depth. An experiment is carried out with objects whose texture depths span a large spectrum that contains most real-life objects. A custom-designed extension of the standard convolutional neural network structure is detailed that helps overcome local variations with multiple input channels. The results show that the system's estimates fit to the texture depth level to a promising degree and a clear trend that increased data set size as well as more complex networks can deliver even better results. The next step in our research is to combine our existing projects of texture estimation with a multi-channel neural network input to deliver the prototype of a fully robust contactless thermographic material characterization system. We will also be looking into how the estimated texture depth correlates to human perception of texture roughness over a range of everyday objects.

References

[1] X. Maldague, *Theory and Practice of Infrared Technology for Nondestructive Testing*, Wiley: New York, NY, USA, 2001.

[2] V. Yefremenko, E. Gordiyenko, G. Shustakova, Y. Fomenko, A. Datesman, G. Wang, J. Pearson, E. E. W. Cohen, and V. Novosad, *A broadband imaging system for research applications*, Review of Scientific Instruments, vol. 80, Article 056104, pp. 1-3, 2009.

[3] R. Usamentiaga, P. Venegas, J. Guerediaga, L. Vega, J. Molleda, and F. G. Bulnes, *Infrared Thermography for Temperature Measurement and Non-Destructive Testing*, Sensors, vol 14, 2014, pp. 12305-12348.

[4] T. S. Durrani, A. Rauf, F. Lotti, and S. Baronti, *Thermal Imaging Techniques For The Non Destructive Inspection of Composite Materials In Real Time*, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 87), 1987, pp. 598-601.

[5] T. S. Durrani, A. Rauf, K. Boyle, and F. Lotti, *Reconstruction Techniques For The Inspection of Composite Materials Using Thermal Images*, International Conference on Acoustics, Speech and Signal Processing (ICASSP 88), 1988, pp. 863-866.

[6] T. Kakuda, A. Limarga, A. Vaidya, A. Kulkarni, and T. D. Bennett, *Non-destructive thermal property measurement of an APS TBC on an intact turbine blade*, Surface & Coating Technology, vol. 205, no. 2, 2010, pp. 446-451.

[7] P. Bison, F. Cernushi, and S. Capelli, *A thermographic technique for the simultaneous estimation of in-plane and in-depth thermal diffusivities of TBCs*, Surface & Coatings Technology, vol. 205, no. 10, 2011, pp. 3128-3133.

[8] F. Cernushi, P. Bison, S. Marinetti, and E. Campagnoli, *Thermal diffusivity measurement by thermographic technique for the non-destructive integrity assessment of TBCs coupons*, Surface & Coatings Technology, vol. 205, no. 2, 2010, pp. 498-505.

[9] S. E. Burrows, S. Dixon, S. G. Pickering, T. Li, and D. P. Almond, *Thermographic detection of surface breaking defects using a scanning laser source*, NDT & E International, vol. 44, no. 7, 2011, pp. 589-596.

[10] I. Plotog, B. Mihailescu, I. Pencea, M. Branzei, P. Svasta, T. Cucu, and M. Tarcolea, *Methods for Pads Thermophysical Parameters Assessment in Terms of 4P Soldering Model*, IEEE 34th International Spring Seminar on Electronics Technology, 2011, pp. 320-326.

[11] N. Horny, J.-F. Henry, S. Offerman, C. Bissieux, and J. L. Beaudoin *Photothermal infrared thermography applied to the identification of thin layer thermophysical properties*, https://www.researchgate.net/publication/265667818, Retrieved October 6th, 2016.

[12] N. W. Pech-May, A. Oleaga, A. Mendioroz, and A. Salazar, *Fast Characterization of the Width of Vertical Cracks Using Pulsed Laser Spot Infrared Thermography*, Journal for Nondestructive Evaluation, vol. 35, article 22, 2016.

[13] S. E. Burrows, A. Rashed, D. P. Almond, and S. Dixon, *Combined laser spot imaging thermography and ultrasonic measurements for crack detection*, Nondestructive Testing and Evaluation, vol. 22, 2007, pp. 217-227.

[14] J.-L. Bodnar, J.-L. Nicolas, K. Mouhoubi, and V. Detalle, *Stimulated infrared thermography applied to thermophysical characterization of cultural heritage mural paintings*, The European Physical Journal Applied Physics, vol. 60, article 21003, 2012, pp. 1-6.

[15] M. Kamel, J.-L. Bodnar, V. Detalle and J.-M. Vallet, *Stimulated infrared thermography applied to the local characterization of fresco*, Quantitative Infrared Thermography Conference (QIRT 16), 2016, pp. 135-143.

[16] K. Mouhoubi, J.-L. Bodnar, V. Detalle, and J.-M. Vallet, *Non-destructive testing of works of art by stimulated by infrared thermography: PPT interest*, Quantitative Infrared Thermography Conference (QIRT 16), 2016, pp. 144-151.

[17] J.-L. Bodnar, J.-L. Nicolas, K. Mouhoubi, J. C. Candore, and V. Detalle, *Characterization of an Inclusion of Plastazote Located in an Academic Fresco by Photothermal Thermography*, International Journal of Thermophysics, vol. 34, no. 8-9, 2013, pp. 1633-1637.

[18] S. Huth, O. Breitenstein, A. Huber, D. Dantz, U. Lambert, and F. Altmann, *Lock-in IR-Thermography a novel tool for material and device characterization*, Diffusion And Defect Data Part B Solid State Phenomena, vols. 82-84, 2002, pp. 741-746.

[19] P. E. Raad, P. L. Komarov, and M. G. Burzo, *Non-Contact Surface Temperature Measurements Coupled with Ultrafast Real-Time Computation*, Twenty-Third Annual IEEE Semiconductor Thermal Measurement and Management Symposium, 2007, pp. 57-63.

[20] C. Ionescu, M. Branzei, B. Mihailescu, and D. Bonfert, *Studies on Thermal Properties of Substrates for Electronics using IR Thermography*, IEEE 20th International Symposium for Design and Technology in Electronic Packaging (SIITME), 2014, pp. 45-49.

[21] *Heat Conduction Equation*, Wolfram Mathworld, http://mathworld.wolfram.com/HeatConductionEquation.html Retrieved Mar. 2, 2019.

[22] J.-C. Krapez, L. Spagnolo, M. Friess, H.-P. Maier, and G. Neuer, *Measurement of in-plane diffusivity in non-homogeneours slabs by applying flash thermography*, International Journal of Thermal Sciences, vol. 43, no. 10, 2004, pp. 967-977.

[23] F. Lakestani, A. Salemo, and A. Volcan, *Modulated spot heating for the measurement of thermal diffusivity*, Journal of Applied Physics, vol. 97, article 013704, 2005, pp. 1-5.

[24] H. Dong, B. Zheng, and F. Chen, *Infrared sequence transformation technique for in situ measurement of thermal diffusivity and monitoring of thermal diffusion*, Infrared Physics & Technology, vol. 73, 2015, pp. 130-140.

[25] N. W. Pech-May, N. Wilbur, A. Mendioroz, and A. Salazar, *Simultaneous measurement of the in-plane and in-depth thermal diffusivity of solids using pulsed infrared thermography with focused illumination*, NDT & E International, vol. 77, 2016, pp. 28-34.

[26] S. N. Pandya, B. J. Peterson, R. Sano, K. Mukai, E. A. Drapiko, A. G. Alekseyev, T. Akiyama, M. Itomi, and T. Watanabe, *Calibration of a thin metal foil for infrared imaging video bolometer to estimate the spatial variation of thermal diffusivity using a photo-thermal technique*, Review of Scientific Instruments, vol. 85, article 054902, 2014, pp. 1-9.

[27] T. Gfroerer, R. Phillips, and P. Rossi, *Thermal diffusivity imaging*, American Journal of Physics, vol. 83, 2015, pp. 923-927.

[28] L. Yeshurun and H. Azhari, *Non-invasive Measurement of Thermal Diffusivity Using High-Intensity Focused Ultrasound and Through-Transmission Ultrasonic Imaging*, Ultrasound in Medicine & Biology, vol. 42, no. 1, 2016, pp. 243-256.

[29] T. Aujeszky, G. Korres and M. Eid, *Measurement-Based Thermal Modeling Using Laser Thermography*, in IEEE Transactions on Instrumentation and Measurement.

[30] T. Aujeszky, G. Korres and M. Eid, *Material Classification with Laser Thermography and Machine Learning*, Quantitative InfraRed Thermography Journal. DOI: 10.1080/17686733.2018.1539895

[31] K. Simonyan and A. Zisserman, *Very Deep Convolutional Networks for Large-Scale Image Recognition*, 2015 International Conference on Learning Representations (ICLR 2015), https://arxiv.org/pdf/1409.1556.pdf Retrieved Feb. 25, 2019.