SpectroPhone: Enabling Material Surface Sensing with Rear Camera and Flashlight LEDs

Maximilian Schrapel  
 HCI Group, Leibniz University
Hannover
Hannover, Germany

Phillipp Etgeton  
 HCI Group, Leibniz University
Hannover
Hannover, Germany

Michael Rohs  
 HCI Group, Leibniz University
Hannover
Hannover, Germany

ABSTRACT

We present SpectroPhone, a surface material sensing approach based on the rear camera of a smartphone and external white LED light sources. Warm and cool white LEDs, as used for dual or quad flashlights in smartphones, differ in their spectral distribution in the red and blue range. Warm and cool white LEDs in combination can produce a characteristic spectral response curve, when their light is reflected from a surface. We show that with warm and cool white LEDs and the rear-camera of a smartphone 30 different materials can be distinguished with an accuracy of 99%. Based on a dataset consisting of 13500 images of material surfaces taken at different LED light intensities, we report recognition rates of support vector machines with different parameters.

CCS CONCEPTS

- Human-centered computing → Displays and imagers.

KEYWORDS

Material sensing, Pattern recognition, Mobile Interaction.

ACM Reference Format:


1 INTRODUCTION

Today, mobile phones are highly integrated systems with a variety of sensors for different applications. For instance, inertial measurement units can track the position of the phone for applications such as barcode readers and games, or pedometers. In addition, rear cameras and flash light LEDs can not only illuminate a scene but also measure the heartbeat [14], blood pressure [17] and Hemoglobin [18] through the skin of a finger. Those and many more applications indicate that the phone is a ubiquitous multi-functional toolbox [5].

Previous research has shown that a front camera in combination with different display color illuminations allows to capture spectroscopic features of surfaces [20]. During the sensing process, the display can not be used for other purposes, e.g. to give feedback on the progress. In this work, we show that the rear camera and white flashlight LEDs are sufficient to capture spectroscopic features. Since white light covers the entire visible spectrum, the number of captured images can be reduced. Due to API limitations, we use external LEDs mounted in a 3D-printed phone case to test different illumination levels.

2 RELATED WORK

Material sensing with smartphones based on spectroscopic features has been performed in the past with different methods. The methods can be divided into approaches that require additional hardware [10, 15, 19] and those that use only built-in capabilities [16, 20] or even embed a molecular spectrometer [2]. All rely on the same principle: Emitted light is reflected by a material and then captured by a sensor. The remaining light is spectrally decomposed, processed and compared to a dataset. Various light sources have been examined in the past: RGB LEDs, also combined with infrared and ultraviolet LEDs or lasers [15]. Without any additional equipment the smartphone display can also be used as a light source and the front camera is then utilized as a sensor [16, 20]. Moreover, front cameras and flashlights have been used together with additional equipment for colorimetric biosensing applications [19]. Additionally, the company Two-Photon research is exploring the opportunity of using smartphone front cameras for instant COVID-19 tests [3].

In addition to cameras, inexpensive light sensors can also detect spectral components of a material [10]. It has been found that expensive spectrometers do not provide good results in comparison to mobile applications because they measure the light only at a small spot and do not detect any variations in the surface texture [20]. Dedicated sensing equipment has been used for robots to sense object’s fragility [8]. Besides spectroscopic material sensing there has also been research on measuring object resonances with vibrations [6, 12] and varying slicing parameters to distinguish different 3D printed objects [7].

In order to distinguish from related work, our goal is to explore whether flashlight LEDs and rear cameras can achieve comparable results to SpeCam [20]. Keeping the display visible would allow further interactions like showing the sensing progress or giving feedback on interfering light. Due to restrictions in controlling the intensity of the flashlight LEDs of mobile phones [4, 9] without hacking the operating system, we evaluated our approach with
Figure 1: Rear view of the prototype. The case of the used smartphone (Huawei P10 lite) contains an extra control unit for the warm (see marker a) and cool white (marker b) LED via Bluetooth. A small frame made of black foam around the sensing window (marker c) prevents ambient light from entering and creates a fixed distance of 3 mm to the target object (marker d). The black housing color has the lowest influence on the measured spectra and is 3D-printed from ABS material (marker e). The battery powered (marker f) PCB for controlling the LEDs (marker i) is shown in the left corner. Bluetooth commands (marker g) are processed by an ATmega microcontroller (marker h) to generate PWM signals.

external white LEDs (similar to the built-in ones) at a range of intensity levels.

3 SPECTROSCOPY WITH WHITE LEDS
Modern smartphones use dual or quad LED flashlights with different color temperatures to illuminate a scene. Their purpose is to achieve more natural colors in the resulting picture depending on the ambient light conditions. For example, in a relatively dark location where warm white light sources with a color temperature between 2000 and 4000 Kelvin are used, cool white LED flashlights with a color temperature of about 5500 to 6500 Kelvin produce an unnatural image. Therefore, warm white flashlights are applied to achieve an adequate illumination with lower blue and higher red components, as can be seen in Figure 2. Furthermore, modern smartphones use flashlights with a high color rendering index (CRI) of over 80 which describes the faithful reproduction of colors [1].

Based on these fundamentals the full visible light spectrum can be reproduced by dual LED flashlights. In order to overcome differences in the hardware configuration of different manufacturers, a simple calibration on a white surface, e.g., a piece of paper, may be performed once. The spectral response can then be adapted to a comparative dataset.

4 METHODOLOGY
Our additional hardware unit inside the housing consists of a Bluetooth transmitter that is connected to a microcontroller that controls the external LEDs by pulse width modulation (PWM) via a current source. We use 700 mA OSRAM OsLon LEDs with a high CRI value of 90 to ensure accurate color reproduction. To avoid streaking (vertical stripes) in the resulting image, the PWM frequency must be significantly higher than the refresh rate of the smartphone’s camera. The ATmega 8 microcontroller supports a PWM base frequency of 31.372 kHz which avoids any vertical streaks in the captured image.

The smartphone application sends commands via Bluetooth to the microcontroller where the PWM signals are generated to control the flashlight intensity. Each image is captured with the maximal resolution of 3986 x 2976 pixels with no auto-focus and ISO 800. For a full sample, 15 images have to be taken based on the LED intensity level. We control each LED in 20 % intensity steps resulting in 5 images per LED. Then we repeat the procedure by controlling both LEDs simultaneously, which results in 5 additional images. All pixels of each image are then broken down into their RGB components and divided into histograms with 64 bins per color channel as

![Normalized spectral composition](image)

Figure 2: Normalized spectral composition for 2700 K (orange) and 6500 K (blue) light sources [11]. It can be seen that cool white LEDs apply more spectral power on the blue component while warm white LEDs focus on the red component. For our studies we used two OSRAM OsLon LEDs with the same distributions.
described in [20]. After a normalization the histograms based on the maximum value, all histogram bins are concatenated to a feature vector. Per image 64×3 histogram bins are generated, which results in a maximum of 2880 feature components for 15 images. The resulting vector is then transmitted to a server to be classified by an Scikit-Learn Support Vector Machine in Python 3 [13]. Our server is a conventional PC using Ubuntu OS for running a Python script to receive feature vectors, classify spectroscopic features, and send the classification result back to the smartphone.

5 DATASET & STUDY
In order to evaluate and compare our approach with related work, we have chosen 30 similar materials as those selected by Yeo et al. for SpeCam [20]. In summary for each sample per class 15 images were taken, as described in Section 4, resulting in 450 pictures for 30 samples. In total for 30 classes 13500 images are contributing to our dataset with a resolution of 12 megapixels. Each sample was taken at a different position and alignment on each material as described by [20]. The dataset together with the casing and PCB layout is available at https://github.com/M-Schrapel/SpectroPhone.

We aimed to evaluate whether the rear camera of a smartphone for image capture and warm and cool white LEDs as the light source are able to achieve similar recognition rates as related work with the front camera for image capture and the front display as the light source [16, 20]. For this purpose we chose three flashlight configurations: 2700 K only, 6500 K only, and both light sources together. For each configuration we evaluated different light modes: single mode with 100 % light intensity, dual mode with 20 % and 100 % intensity, and full mode with all contributing images per sample. When only one LED is controlled, 5 images are taken in full mode. With both LEDs in full mode 15 images are contributing to a sample. For our dataset all samples were taken in full mode with both LEDs. Other configurations can be derived by selecting the corresponding features. The goal of our tests is to achieve a trade-off between the number of images and the recognition rate while minimizing the time to capture. The maximum feature vector length when both LEDs and full mode are applied is 2880 components resulting from 64 histogram bins for 3 color channels and 15 images per sample. In single mode with one LED, a minimum length of 192 components is used. If both LEDs are controlled in dual mode, each single LED is controlled individually and then also together with 20 and 100 % intensity resulting in a feature vector with 1152 components.

The SVM used for classification is also examined more closely. It has to be determined whether the selected materials can be distinguished by means of a simple dividing hyperplane. Decisive for this are the kernel function and the size of the margin of the decision boundary, which is set via the penalty parameter C. We use 10-fold cross-validation with a split ratio of 70:30 to evaluate our models.

6 RESULTS
To estimate the performance of our approach and find the most appropriate SVM parameters, GridSearchCV [13] is applied on all LED configurations and SVMs using linear, rbf and polynomial kernels. We tested C parameters from 0.001 to 10000 for our analysis. The best results for each LED configuration and light mode are listed in Table 1.

In comparison to related work, white LED flashlights and the rear camera also provide high recognition rates of about 99 % which is similar to the results of SpeCam [20].

The linear and rbf SVM kernels show the best results for the most configurations. Beside of the entries in Table 1, where the rbf kernel had the highest accuracy, the linear kernel always was close, as exemplified for both LEDs in full mode. By further evaluating the applied SVM parameters found by GridSearchCV [13] it should be mentioned that the chosen features tend to be linearly separable by a decision boundary. For instance the polynomial kernel has the highest recognition rate when the degree of the equation is set to 1, which is almost equivalent to the linear kernel. The best performance was achieved when both LEDs contribute one image at full light intensity to the feature vector. While SpeCam [20] uses 7 images, flashlight LEDs can reduce the time for material sensing since only two images are sufficient for comparable recognition rates.

A further evaluation of the confusion matrix shows which classes are confused with each other despite an optimal choice of parameters. For this purpose in Figure 4 the linear SVM with both LEDs and all 15 images per sample was chosen. The only remaining confusion

![Figure 3: Subset of used materials and corresponding histograms with different LED configurations. The histograms differ clearly from each other. The classes BookGreen and BookRed show high proportions on another color at low color values due to the additive color mixing. In addition, it is noticeable that the resulting histograms of the warm white LED are more similar to the spectrum of both LEDs.](image)

Table 1: Predictive accuracy of the best results with a 10-fold cross-validation.

<table>
<thead>
<tr>
<th>LED</th>
<th>Intensity mode</th>
<th>SVM Kernel</th>
<th>SVM C Parameter</th>
<th>Accuracy</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>4500K</td>
<td>single</td>
<td>rbf</td>
<td>1000</td>
<td>0.950</td>
<td>0.0181</td>
</tr>
<tr>
<td>4500K</td>
<td>dual</td>
<td>linear</td>
<td>60</td>
<td>0.948</td>
<td>0.0172</td>
</tr>
<tr>
<td>4500K</td>
<td>full</td>
<td>linear</td>
<td>60</td>
<td>0.951</td>
<td>0.0187</td>
</tr>
<tr>
<td>3700K</td>
<td>single</td>
<td>rbf</td>
<td>1000</td>
<td>0.967</td>
<td>0.0097</td>
</tr>
<tr>
<td>3700K</td>
<td>dual</td>
<td>linear</td>
<td>60</td>
<td>0.966</td>
<td>0.0122</td>
</tr>
<tr>
<td>3700K</td>
<td>full</td>
<td>linear</td>
<td>60</td>
<td>0.965</td>
<td>0.0090</td>
</tr>
<tr>
<td>bcool</td>
<td>single</td>
<td>rbf</td>
<td>60</td>
<td>0.969</td>
<td>0.0084</td>
</tr>
<tr>
<td>bcool</td>
<td>dual</td>
<td>linear</td>
<td>60</td>
<td>0.968</td>
<td>0.0105</td>
</tr>
<tr>
<td>bcool</td>
<td>single</td>
<td>polynomial</td>
<td>60</td>
<td>0.968</td>
<td>0.0092</td>
</tr>
<tr>
<td>bcool</td>
<td>full</td>
<td>rbf</td>
<td>10000</td>
<td>0.969</td>
<td>0.0099</td>
</tr>
<tr>
<td>bcool</td>
<td>full</td>
<td>linear</td>
<td>60</td>
<td>0.968</td>
<td>0.0105</td>
</tr>
</tbody>
</table>

![Figure 4: Linear SVM with both LEDs and 15 images per sample.](image)
the detection of counterfeit clothing is possible on the basis of the
used fabrics. A small sensing area on the clothing could eliminate
the need for additional shieldings around the camera and flash-
light. A user study could then identify whether our approach is
a more natural interaction technique than using the display for
material sensing. It would also be of interest to identify whether
potential users would trust such a system and tend to avoid coun-
terfeit clothing when using this technology. Furthermore, we are
going to examine the generalizability of the approach to differ-
ent smartphone models and their built-in flashlights, provided API
limitations are removed.

7 CONCLUSION & DISCUSSION

This work reported the use of built-in hardware of smartphones,
namely the rear camera and white flashlight LEDs, for material
sensing. Based on the measured RGB values at different intensity
levels and LED combinations, high recognition rates of approxi-
mately 99 % were achieved on a dataset of 13500 images, consisting
of 30 materials with 30 samples of each material and each sample
consisting of 15 images taken at different illuminations. Due to API
limitations in controlling the intensity of flashlight LEDs of current
mobile phones [4, 9], external LEDs, providing the same spectral
distribution as built-in LEDs, were connected to a microcontroller
and Bluetooth unit and placed in a housing.

The results show that only a few images per surface are sufficient
to achieve acceptable recognition rates for typical surfaces. How-
ever, black surfaces turn out to be difficult to recognize with at least
89%. We had also various white surfaces in our dataset (Cardboard-
White, Plate Porcelain, ThieleWhite, Paper, Wall White, Whiteboard),
which proves that our proposed approach is not only based on
color separation. The results in Table 1 indicate that built-in dual
flashlights can reduce the number of images to be taken for mate-
rial sensing. With two images (warm and cold white illumination)
at 100 % intensity comparable recognition rates to SpeCam [20]
were achieved. With our approach the time for the material sensing
process can be reduced. Furthermore, we could half the number of
calculated features from 768 [20] to 384. If the API restrictions were
removed, the test could be repeated with the embedded flashlight
and the application could be made available to a wide range of
smartphone models. The case could then be replaced by a shielding
around the camera and flashlight. The opportunity of keeping the
display visible enables to show the sensing process as well as the
results instantly. Feedback could be given when external light in-
terferes with the sensing process similar to camera-based heartbeat
measurements [14].

In the future, we are going to extend the presented approach to enable
material recognition while simultaneously using the display.
As one example application, we are going to investigate whether

Figure 4: Confusion matrix for a linear SVM with both LEDs
and all 15 images per sample. Only black surfaces show con-
fusion, because for them the absorption of light is strongest.

REFERENCES

[5] Rafael Ballagas, Jan Borchers, Michael Rolfs, and Jennifer G. Sheridan. 2006. The
Smart Phone: A Ubiquitous Input Device. IEEE Pervasive Computing 5, 1 (Jan.
2006), 70–77. https://doi.org/10.1109/MPRV.2006.18
put: Back-EMF in Linear Resonant Actuators to Enable Touch, Pressure and
on User Interface Software and Technology (Virtual Event, USA) (UIST '20). Asso-
doi.org/10.1145/3379337.3415823
[7] Mustafa Dogan Dogan, Faraz Faruqi, Andrew Day Churchill, Kenneth Friedman,
on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI ’20). Associa-
3313831.3376202
(2019), 790–797. https://doi.org/10.1109/LRA.2019.2895293
Accessed: 2021-02-02.
Symposium on User Interface Software and Technology (Monterey, CA, USA) (UIST
ham, and W Chan. 2016. Saving energy with light? Experimental studies assessing
the impact of colour temperature on thermal comfort. Energy Research & Social
2019. ViEye: Vibration-Mediated Object Recognition for Tangible Interactive
Applications. In Proceedings of the 2019 CHI Conference on Human Factors in
Computing Systems (Glasgow, Scotland Uk) (CHI ’19). Association for Comput-
ing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300906
Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cour-
Heart Beat Rate using a mobile phone. In 2010 Annual International Conference of
the IEEE Engineering in Medicine and Biology.
[15] Munehiko Sato, Shigeyoshi Yamada, Alex Olwal, Boxin Shi, Atsushi Hiyama, Tomo-
hiro Tanaka, Michitaka Hirose, and Ramesh Raskar. 2015. SpecTrans: Versatile
Material Classification for Interaction with Textureless, Specular and Transparent
Surfaces. In Proceedings of the 33rd Annual ACM Conference on Human Factors in
Computing Systems (Seoul, Republic of Korea) (CHI ’15). ACM, New York, NY,
USA, 2191–2200. https://doi.org/10.1145/2702123.2702169
geschaftsbereiche/biowissenschaften/forschung.html#tabpanel-25. Ac-
cessed: 2021-02-02.
[17] Edward Jay Wang, Junyi Zhu, Mohit Jain, Tien-Jui Lee, Elliot Saba, Lama Nach-


