

## Towards High-Throughput Autonomous Infrared Spectromicroscopy

Petrus Zwart<sup>1,2,3</sup>, Liang Chen<sup>1,2</sup>, Marcus Noack<sup>2,4</sup>, Steven Lee<sup>1,3</sup>, Patricia Valdespino Castillo<sup>1,3</sup>, Hoi-Ying Holman<sup>1,3</sup>

<sup>1</sup>Molecular biophysics and Integrated Bioimaging Division, Lawrence Berkeley Laboratory

<sup>2</sup>Center for Advanced Mathematics in Energy Research Applications, Lawrence Berkeley Laboratory

<sup>3</sup>Berkeley Synchrotron Infrared Structural Biology Program, Lawrence Berkeley Laboratory

<sup>4</sup>Computational Research Division, Lawrence Berkeley Laboratory

phzwart@lbl.gov

Infrared (IR) absorption spectromicroscopy is a powerful, non-invasive probe that provides access to spatio-chemical information at the micron scale. The physical basis of IR spectroscopy lies in the oscillations of dynamic dipole moments in chemical bonds, with resonant frequencies in the IR spectral region of 4,000-400  $\text{cm}^{-1}$  wave numbers. The bending or stretching of chemical bonds between atoms with different electronegativities, such as O-H or C=O, will lead to intense absorption and thus provide a unique fingerprint of specific chemical groups within the sample. The presence - or absence - of specific spectral fingerprints provide the opportunity to locate and identify chemical processes throughout the sample and track these processes in time or as a function of external perturbation.

We can perform these types of measurements using a Synchrotron Fourier Transform Infrared (SFTIR) spectromicroscopy setup, as it provides orders of magnitude more photons than traditional bench-top machines<sup>1</sup>. Even with an ultra-bright IR source as provided at the Advanced Light Source, acquisition times typically take multiple hours. These long acquisition times are in part caused by the size of the field of view, as compared to the probe size: users typically analyze a 70  $\mu\text{m}$  by 100  $\mu\text{m}$  sample using a regular grid with a spacing of 1  $\mu\text{m}$ . With an acquisition time of 4 seconds per pixel, we would need about 8 hours to measure the full sample. Given that access to instruments is scarce, compounded by the desire to characterize different samples, being able to speed up data acquisition is of paramount importance.

Here we present a strategy that drastically increases the efficiency of SFTIR spectromicroscopy by coupling the data collection with Gaussian Process based surrogate model<sup>2,3</sup>. This approach models the full hyperspectral datasets across the entire field of view, including regions we haven't measured yet, by multivariate Normal distribution. By analyzing this distribution, we gain insight into which future measurement locations provide the greatest reduction in total uncertainty and can also predict various quality metrics of the surrogate model we can use to end an experiment. Preliminary experiments show we can increase the throughput of SFTIR experiments by a factor of  $\sim 20$ .

1. Holman, H. -Y N. & Martin, M. C. Synchrotron Radiation Infrared Spectromicroscopy: A Noninvasive Chemical Probe for Monitoring Biogeochemical Processes. *Advances in Agronomy* 79–127 (2006) doi:10.1016/s0065-2113(06)90003-0.
2. Chang, H. *et al.* Building Mathematics, Algorithms, and Software for Experimental Facilities. in *Handbook on Big Data and Machine Learning in the Physical Sciences* 189–240 (World Scientific, 2020).
3. Noack, M. & Zwart, P. Computational Strategies to Increase Efficiency of Gaussian-Process-Driven Autonomous Experiments. in *2019 IEEE/ACM 1st Annual Workshop on Large-scale Experiment-in-the-Loop Computing (XLOOP)* 1–7 (2019).

**Keywords:** Autonomous experimentation