Computer Vision Techniques for the Energy Calibration of the Hard X-Ray Single-Shot Spectrometer at the European XFEL

PREPARED FOR

MD Zishan

PREPARED BY

Christian Grech

MXL - DESY

1. Project Overview

Single-crystal monochromators are used in free electron lasers for hard x-ray self-seeding, selecting a very narrow spectral range and further amplifying the original SASE signal. When a crystal is rotated, one can exploit several symmetric and asymmetric reflections as established by Bragg's law. Currently, a Machine Learning (ML) classifier is used during experimental setup to identify the crystal indices corresponding to a given reflection, compare to a model, and eventually calculate the difference between the photon energy as measured by the single-shot spectrometer and the actual one. Extracting information from images has proven to be challenging measurement artifacts such as noise and dead pixels.

This project proposes an alternative method for identifying the actual photon energy by implementing a technique known as **Template Matching**. Template matching is a common computer vision technique where an algorithm is trying to find similarities between two images. This technique slides a window across the image that will provide a percent match with the template. If the percent match is above a certain threshold then it is assumed to be a match. This will in some situations result in poor performance because it requires the template to **overlap the image identically**.

Template matching is a sufficient technique if one knows exactly what they are looking for and are confident that it will appear almost identically in every example. ML techniques such as Convolutional Neural Networks (CNN) can be used in cases where the template and the image are not identical. ML techniques are not rigid and are able to generalize a dataset very well. What is required in this case is a fairly large dataset so that the model is robust to many different scenarios.

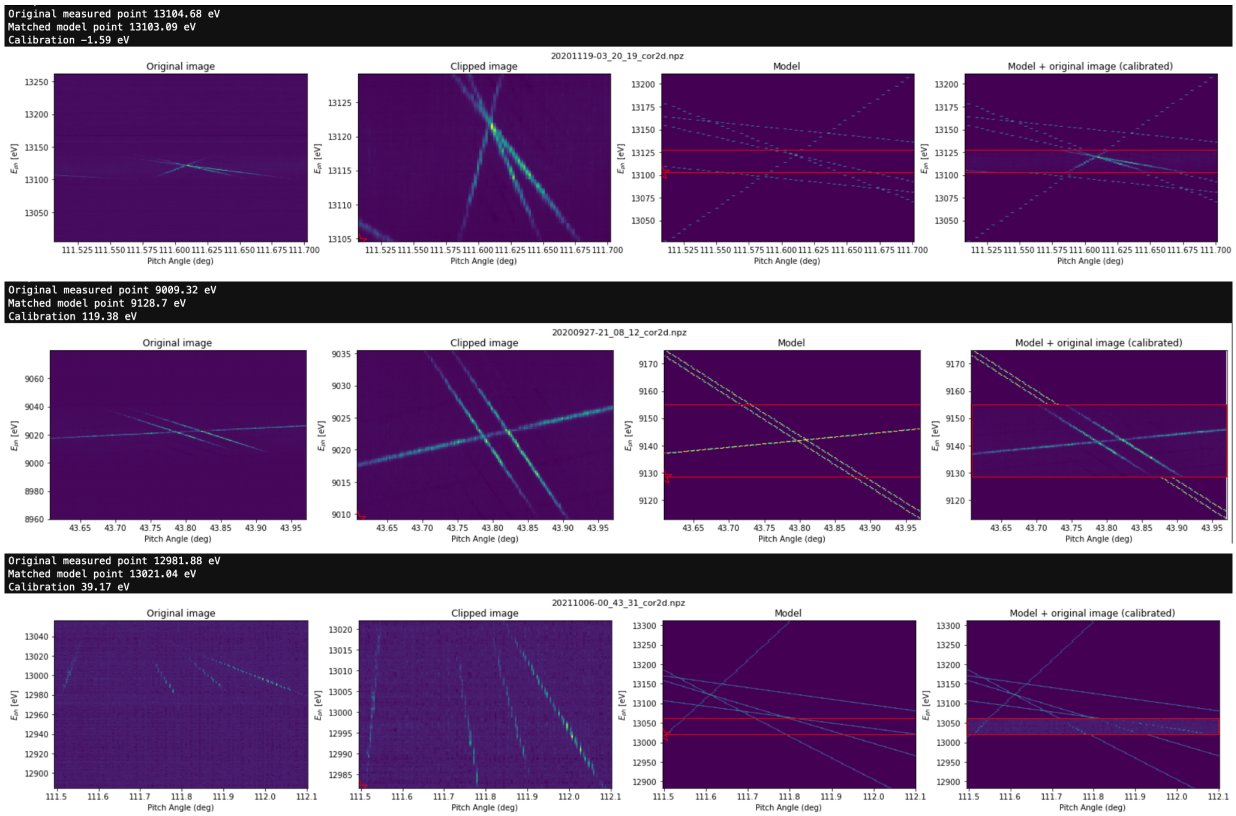
This project involves the matching of a **processed spectrometer image (template)**, with an image created from an available model. The student will have the opportunity to work with experimental data from the European XFEL. Moreover, the student will obtain a general overview of several image processing/ML techniques and will be able to carry such knowledge into their future studies.

2. Obstacles

- The project involves the manipulation and processing of spectrometer image scans. Some scans are **noisy** and **faint** which present a challenge.
- The aim is to create images from a model which is available and compare them with the measured scans. Hence the created images must **resemble the measured scans** as much as possible for more accurate results.
- Packages for template matching exist already. Are these packages **consistent** and **resilient** to possible variations between measurements and model? What happens when scaling, rotation transformations are present?
- If the use of packages is not sufficient, maybe the use of **Convolutional Neural Networks** can help? The resources below show some examples where this was implemented.

3. Proof-of-concept

Three examples from the available spectrometer images are shown below. The following proof-of-concept gets the original spectrometer scan (first image) and clips out the unnecessary blank areas (second image). The second image is the template. The third image is the image created from the model, with the red box being the matched area using sklearn's *match_template* package. By comparing the bottom left corner energy value of the spectrometer image and the matched image, one can calculate the energy difference. The final image plots the calibrated template on top of the model.



4. Software

Possible packages:

- Python
- Sklearn
- PyTorch
- Keras
- OpenCV

The proof-of-concept code will be provided as well as a dataset of around 150 images. Images can also be artificially created to enlarge the dataset.

5. Milestones

Milestone	Tasks
1 - Research (1 week)	
1.1	Background reading on the current status of the project.
1.2	Familiarisation with spectrometer images and manipulation.
1.3	Reading on Template Matching and projects implemented in this area.
2 - Development (4 weeks)	
2.1	Image cleaning and processing (both spectrometer images and artificially created images from model)
2.2	Implementation of Template Matching using standard packages
2.3	In the case a model is required: creating dataset with images and targets
2.4	Neural network training and evaluation
2.5	Neural network hyperparameter optimization.
3 - Performance evaluation (1 week)	
3.1	Comparison of energy calibration calculations with the current dataset.
3.2	Evaluate methods to improve performance, efficiency.
4 - Comparison with ML classifier (1 week)	
4.1	Compare accuracy, reliability against the current ML classifier method.
5 - Reporting (1 week)	
5.1	Finalise documentation.

6. Resources

- https://mattmaulion.medium.com/template-matching-image-processing-6eb1d5425248
- https://scikit-image.org/docs/dev/auto_examples/features_detection/plot_template.html#template-matching
- http://cs231n.stanford.edu/reports/2017/pdfs/817.pdf
- https://journals.iucr.org/j/issues/2019/04/00/fs5173/fs5173.pdf
- https://github.com/XingruiWang/wheres-waldo