

Automatic detection and of dipoles in large area SQUID magnetometry

Lisa Qian

December 14, 2012

1 Introduction

1.1 Scanning SQUID magnetometry

Scanning SQUID magnetometry is a powerful tool for metrology of individual nanomagnets because of its incredible flux sensitivity (as low as 100 electron spins) and non-invasive nature. With micron-sized pick-up loops, this technique has been used to study individual magnetic dipoles in a variety of systems such as magnetotactic bacteria, nanofabricated bar magnets, and naturally occurring nanomagnetic patches in complex oxide heterostructures. Figure 1a shows a cartoon of a scanning SQUID magnetometer over individual point dipoles. A DC magnetometry image of a silicon substrate containing individual magnetotactic bacteria is shown in Figure 1b, where the individual dipole orientations are clearly visible. Such an image is a convolution of the local magnetic field component perpendicular to the pick up loop with the point spread function defined by the pick up loop geometry.

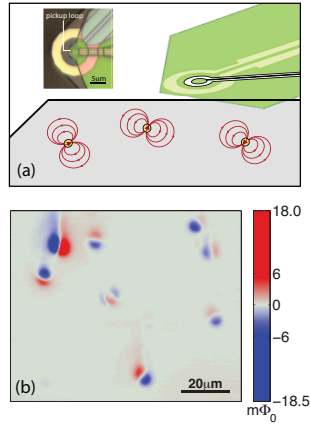


Figure 1: **(a)** Cartoon of a scanning SQUID imaging magnetic dipoles. **(b)** A typical magnetometry image, which is a convolution of the sample's local magnetic field with the SQUID pick up loop shown in (a).

Using both fine and course motion scanners, we can cover a sample area as large as $5\text{mm}\times 5\text{mm}$. Each scan can be as large as $300\mu\text{m}\times 300\mu\text{m}$, and often contain many individual dipoles that may be clustered together. Once a scan is taken, we often need to analyze properties of each individual dipole such as their dipole moments. Current practice is to identify each dipole by eye, manually crop the image, and send the cropped image to an analysis program. In crowded images, selection of dipoles is a slow and monotonous process. In addition to finding each dipole, the crop window, or bounding box, for each dipole is also of importance. Since magnetic fields are long range, we want to make the bounding box as large as possible without picking up magnetic signals from neighboring dipoles or uneven background.

The goal of this project is to automatically detect magnetic dipoles given a large area magnetometry scan and to determine the appropriately sized bounding box.

2 Supervised learning

The idea is to go through the entire large area scan with a sliding box of fixed dimensions. Each image is an example and is classified as either a part of a dipole (1) or not (0). Because stronger dipoles extend across more space than weaker dipoles, many boxes classified as (1) might be needed to cover an entire dipole. At the end, neighboring “1” boxes are merged to form the bounding box for a dipole.

2.1 Challenges

Classification and feature design for this project was challenging for several reasons. First, many magnetometry scans have slowly varying background signals that can be on the order of the dipolar signals themselves. But because the dipolar fields themselves are long range, it is often challenging to properly remove background fields while preserving the structure of the dipoles. Second, the strengths of the dipoles we measure vary across roughly four orders of magnitude (from $10^5\mu_B$ to $10^9\mu_B$). Thus the spatial extent of the dipoles also vary greatly, ranging from approximately $5\mu\text{m}\times 5\mu\text{m}$ for the weakest dipoles to $35\mu\text{m}\times 35\mu\text{m}$ for the strongest ones. Initially I used a very large feature vector to cover all the entire range of pixel intensity, along with sliding boxes of varying sizes. Both of these ideas proved to be not computationally feasible.

2.2 Feature selection and SVM

The background problem was solved by using the gradient of the large area scan instead of the magnetometry image itself. Figure 2 shows a magnetometry image with particularly strong background fields and how simply taking the gradient nearly completely removes the background. To obtain training/test examples, I take the gradient of the large area magnetometry scan and go through this gradient image with a sliding window as described previously.

The feature vector for each image obtained through with the sliding window is a histogram of pixel intensity of that image. I discretize the feature vector linearly from

$-10\text{m}\Phi_0/\mu\text{m}$ to $+10\text{m}\Phi_0/\mu\text{m}$ in increments of $0.1\text{m}\Phi_0/\mu\text{m}$. For a given training/test example, I count up the number of pixels in that example corresponding to each discretization level. This results in a length 200 sparse feature vector. I used an SVM to train and classify.

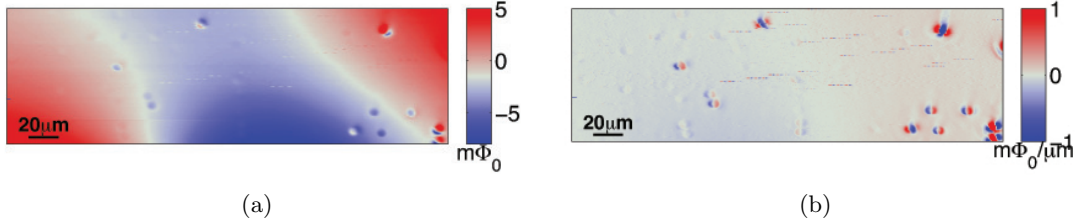


Figure 2: **(a)** Large area magnetometry scan, without background removal. **(b)** The gradient of (a), without additional background filtering.

2.3 Training and test examples

I collected training samples by looking at a total of seven large area scans. These scans contain a total of 193 dipoles. To get a good variety of dipoles, I chose large area scans taken at different heights and with levels of crowdedness. A large scan image is usually around $300\text{ pixels} \times 400\text{ pixels}$ and covers about $300\mu\text{m} \times 300\mu\text{m}$. The size of the sliding window is set to 22×26 pixels, which is the size of the smallest dipole I found amongst the seven large area scans I looked at. For training examples, I skipped by 15 pixels between each image. For test data, I skipped by 10 pixels between images.

It is important to note that since dipolar fields are long range, there is no precise way of determining when a dipole “ends and draw the appropriate bounding box. In my work here, I labeled an example as ‘0’ when any stray dipolar fields are on the order of the noise level of the background.

3 Results

3.1 Leave one out cross-validation

Because I had a limited sample size (seven large area scans), I used LOOCV to train, test and compare my models. Using an SVM with a polynomial kernel gives 81% precision and 83% recall. An SVM with a Gaussian kernel gives 85% precision and 87% recall. Only four false positives were found among all 7 large area scans.

These results are very satisfactory - the dipoles that were missed by the classifier were very weak and right at the noise level. Figure 3 a and c show the results of the model on two large area scans. Each of these scans were classified using an SVM with Gaussian kernel trained on the other six scans. Further improvements could be made by including more positive training data.

3.2 Merging to form bounding box

The final step is to merge neighboring test images that have been classified as ‘1’ (part of a dipole) to form the final bounding box for the dipoles. Figure 3 **(b)** and **(d)** show the final results. While this process works well on scans with sparse dipoles (**(b)**), it does not work well for crowded scans (**(d)**). This is because the sliding window is currently quite large and can cover more than one dipole at a time in crowded images.

While this problem will be alleviated using a smaller window size (breaking up the scan into more and smaller test images), it will always be a problem because of the long range nature of the dipolar fields. When the dipoles are crowded, it becomes impossible to separate out individual dipoles from a cluster. In these cases, I would eventually like to identify the cluster as a separate classification.

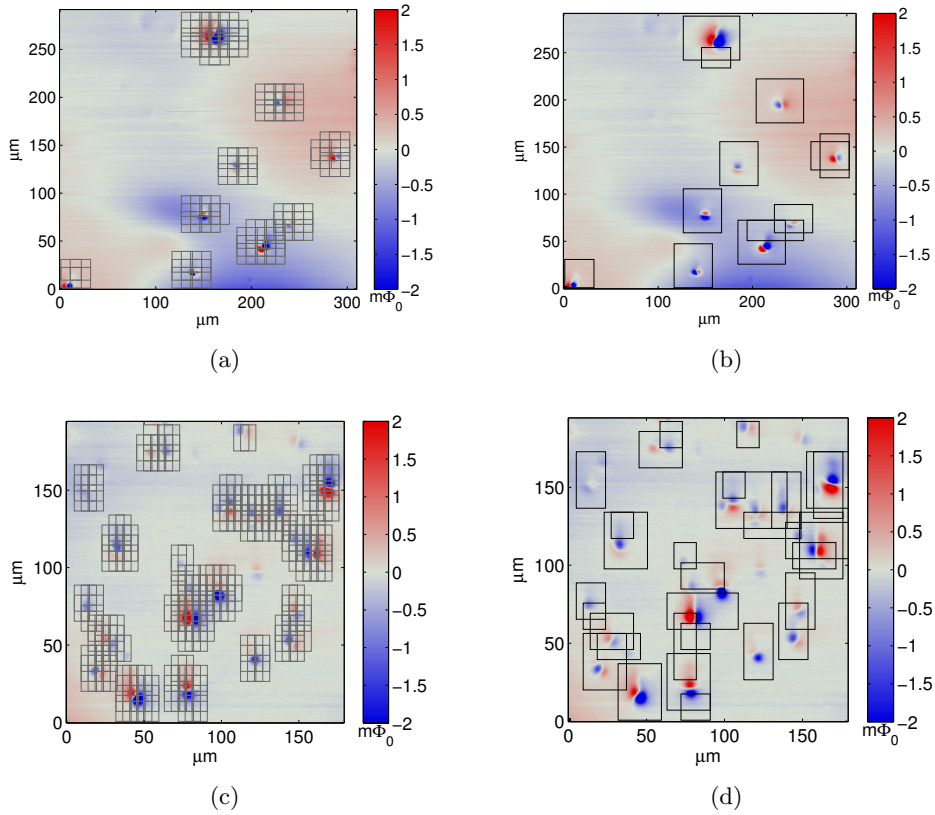


Figure 3: **(a)**Results of SVM classification on a large area magnetometry scan. **(b)**After merging to form one bounding box for each dipole. **(c) - (d)**The same for another, more crowded scan.

3.3 Conclusions

In this project, I successfully trained an SVM classifier to automatically detect dipoles in a large area SQUID magnetometry image. Using a Gaussian kernel yields slightly better results over a polynomial kernel, with a precision of 85% and recall of 87%. These results are satisfactory and can be improved by including more positive training examples and reducing the sliding window size. Significant improvements still need to be made to draw appropriate bounding boxes in crowded images.