

EENG580 Activity 2: Image Registration

Activity summary

Overview: Use Fourier properties to compute shifts between similar images

Setting: Develop in class, then collect data around campus.

Curricular elements: both tinkering and gaming

Prerequisites: basic familiarity with MATLAB

Topics/concepts covered: Fourier transform properties, basic image representation in MATLAB

Learning outcomes: After completing this activity, students should be able to:

- explain the differences between a Fourier transform of a 1-D and a 2-D signal
- explain how a time/spatial shift affects the frequency domain
- explain how the difference between a linear and a circular shift affects the validity of Fourier properties

Expected time to complete: two 2-hour lab sessions and 4-8 hours of work outside of class

Required hardware/materials: Minimum: a digital camera; typical smart phone cameras are adequate. A tripod is encouraged.

Required instructor interaction: partially supervised, with occasional guidance

Common mistakes/pitfalls: Hand-held cameras will probably produce some rotation between frames; the instructor can provide translation-only test frames by taking two copies of an image and trimming different edges off. Many students use `fft.m` when they should use `fft2.m`. The third dimension in images (indexing the color) confuses some students.

Method of assessment: instructor graded, based on final product; plus peer voting on subjective evaluations of the competitions

Contributors: Richard Martin, Andrew Klein, Cody Watson, and Jennifer Hefner

The views expressed in this document are those of the authors, and do not reflect the official policy or position of the United States Air Force, United States Navy, Department of Defense, or the U.S. Government. This document has been approved for public release; distribution unlimited. This work was supported by ONR grant N000141512442, and is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. <http://creativecommons.org/licenses/by-sa/4.0/> ©2016 by A.G. Klein & R.K. Martin. Contact: andy.klein@wwu.edu and richard.martin@afit.edu.





Figure 1: “Spot the difference” images from [8], used with permission, with slight misalignment added.

1 Introduction and motivation

Image registration involves aligning several different images of the same scene [1], [2], [3], [4]. For example, the images may be taken from a camera mounted in a moving vehicle, from different sensor platforms, or from the same platform but at different times. Image registration has applications in digital watermark detection [5], medical imaging [6], and remote sensing [7]. Some specific applications of registration include noise averaging, in which many frames are averaged to improve the Signal to Noise Ratio (SNR); microscanning (a.k.a. super-resolution), in which the overall resolution is improved by noting that the shifts between frames may be non-integer numbers of pixels; and change detection, in which frames from two successive days are compared. You may explore some of these applications if you participate in the competitions for this project.

Consider the two jpeg images in Fig. 1, obtained from the All Star Puzzles website [8]. Image an automated change detection application wherein you need to find all the differences between the two images. This is a simplified approximation of a situation that may occur in military or law enforcement scenarios, in which it may be important to note minor changes to a street-level environment. However, the two images shown here are slightly misaligned, so a pixel-by-pixel comparison will be thrown off. Thus, before any application can be pursued, the images must be re-aligned, and this process is called “image registration.”

In general, misalignment effects includes rotation, scaling, and translation between images, but at first we will only consider translation. This translation can be modeled as

$$x_2[m, n] = x_1[m - M, n - N], \quad (1)$$

ignoring edge effects. Recall from the lecture material that in one dimension (i.e. time instead of two spatial

dimensions), a translation in the original domain induces a linear phase term in the frequency domain [9]:

$$\begin{aligned}\mathcal{F}\{x[n]\} &= X(\omega), \\ \mathcal{F}\{x[n-N]\} &= X(\omega) \cdot e^{-j\omega N}.\end{aligned}\tag{2}$$

A similar thing will happen in 2D, though there will be two phase terms, one for each spatial frequency. Deriving and exploiting this phenomenon for efficient computation is the basis for this project.

2 Getting started

This is a very open-ended activity. Everyone will have a different approach, or a different “answer.” This is likely to be a very frustrating exercise at times, but that is ok; the instructor is here to help you, and you should not hesitate to ask questions. A lot of information is intentionally missing from this lab description; you will need to ask questions to complete this lab. The goal of the activities in this course is to encourage a “maker” mindset wherein you feel free to tinker and explore different approaches. We want you to fail, rapidly and frequently, because that is how you identify and correct errors. Your goal is to get comfortable with the idea of rapid prototyping and rapid response.

Pick a lab partner for this activity. You will collaborate on the process, but you will write up individual reports and receive separate grades. Make sure that at least one person per group has access to a digital camera; a smart phone camera should be adequate. For the more advanced aspects of this project, it may be beneficial to have access to multiple, disparate cameras; but that is not a strict requirement. Take several photos and figure out how to get them into MATLAB.

Explore the MATLAB commands listed below. Some or all of them may be helpful, particularly the first few.

- `imread`, `imwrite`, and `image`
- `double` and `uint8`
- `fft2` and `ifft2`
- `abs` and `angle`
- `circshift`
- `ones` and `zeros`
- `size`, `max`, and `mod`

Progress to the point where you can take a photo, read it into MATLAB, display it within a figure window, modify it, and save it as a jpeg. If you can accomplish these tasks, you have sufficient understanding of the hardware interface to complete the entire activity.

3 Tasks

For the first week of this activity, you are expected to study the following material in advance: the discrete Fourier transform (DFT), signal building blocks, DTFT/DFT examples (called “transform pairs” in many textbooks), and DTFT/DFT properties. Before the second week, you should study the material on DFT resolution and the Fourier-Mellin transform. Videos covering these concepts are available on the course website, and you are encouraged to read additional background material in any of the suggested reference textbooks.

The core task in this assignment is to derive and exploit the formula for the frequency-domain effects of a two-dimensional translation of an image. You will use this to re-align (or “register”) two slightly misaligned images. To enable this, explore how to address the following issues:

- (i) What does the formula for a two-dimensional DTFT (or DFT) look like? What is the 2D property analogous to (2)? Prove it analytically if you can.
- (ii) What are good choices for all the variables involved? Remember that you have variables for spatial indices, frequency indices, shifts, sizes, and potentially a color index to account for the RGB nature of jpeg images.
- (iii) Consider (2). How can you manipulate some combination of the frequency domain data $X_1(\omega)$ and $X_2(\omega)$ to isolate the “extra” phase term? What will that isolated term look like (theoretically) if you inverse transform it back into the spatial domain?
- (iv) How can you implement the theoretical operations from (iii) in MATLAB? What does the result tell you about the amount of shift? The images shown in Fig. 1 are available on the course webpage for you to use in testing your initial code.

That should be enough to get a baseline image registration algorithm up and running. Once you are comfortable with that, consider the following issues:

- (v) Are there issues with edge effects? ...the difference between linear and circular shifts? ...noise? ...image compression? ...real vs. complex data? ...MATLAB starting its indexing at one instead of zero? ...negative shifts? ...non-integer shifts?
- (vi) What are some ways to deal with the length-3 color dimension? what are the pros and cons of each method you come up with?
- (vii) What physical impairments can cause difficulty for your registration process?

You will probably complete your initial explorations during the first class session. Between the two in-class sessions, use your time to gather and save some photos. Do your best to ensure that there is no rotation, range differences, or perspective differences between your “easy” photos, but you are encouraged to also collect some “hard” photos that exhibit these effects or effects related to the competitions listed below. In the second in-class session, focus on processing the saved data using multiple approaches, and on interacting with the instructor regarding any difficulties in your approach. Then use some more out-of-class time to take your final data sets and write up your report.

4 Deliverables

You will turn in a report written using the LaTeX word processing system. A template is available on the course webpage. The report should conform to the style guide for IEEE Signal Processing Society conferences, such as ICASSP, ICIP, or GlobalSIP. At a minimum, explain your measurement methodology, your data processing methodology, and your results. Both format and content matter – in particular, use correct grammar and spelling, and revise your text as needed to make sure you are explaining yourself coherently. Include code sparingly, if at all; in general, it is better to include an algorithm in pseudo-code as a figure.

You should strive to always explain why you did what you did, not just what you did. Your results should all be repeatable – maybe not exactly, but comparably. That is, another researcher may not have access to the exact same environment as you or record the exact same realization of random noise, but you want to enable them to follow the exact same process as you. Analysis is also highly encouraged – explain why the results are what they are, not just what they are. For this activity, photos or sketches of the geometry may help with that.

To improve your grade, include a more detailed and coherent explanation of your results, try a more clever or innovative approach, and include more data and analysis. Feel free to diverge somewhat from the stated tasks if you think the situation warrants it. “Out of the box” thinking will always be rewarded.

5 Competitions

The written report constitutes 90% of the grade, and the remaining 10% will be based on competitions. Competitions will include best demonstration of various extensions:

- change detection between two frames,
- mosaicking (combining partially overlapped images into one larger image),
- de-jittering video frames,
- dealing with illumination differences,
- accounting for rotation and/or scale differences (hint: look up the Fourier-Mellin transform, a.k.a. the Fourier transform in log-polar coordinates)
- and super-resolution (combining many low-resolution frames to obtain a high-resolution image).

Competitions will be determined by presenting results to the class, immediately before the start of the next activity. Due to potential subjectivity in image analysis, the winners will be determined by class vote. The first place group in each category will receive a 5% bonus (for each group member) and the second place group members will each receive 2.5%; and no individual may receive more than 10% total. Based on typical class sizes, that means about a quarter of you may get about 10% and another quarter may get about 5%.

References

- [1] L. Gottesfeld Brown, “A Survey of Image Registration Techniques,” *ACM Computing Surveys*, vol. 24, pp. 325–376, Dec. 1992.
- [2] B. Zitová and J. Flusser, “Image Registration Methods: a Survey,” *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, Oct. 2003.
- [3] D. Robinson and P. Milanfar, “Fundamental Performance Limits in Image Registration,” *IEEE Trans. on Image Processing*, vol. 13, no. 9, pp. 1185–1199, Sept. 2004.
- [4] A. Bardera, M. Feixas, I. Boada, and M. Sbert, “Compression-based Image Registration,” in *Proc. Int. Symposium on Info. Theory*, Seattle, WA, July 2006, pp. 436–440.
- [5] H. Cheng and M. A. Isnardi, “Spatial Temporal and Histogram Video Registration for Digital Watermark Detection,” in *Proc. Int. Conf. on Image Processing*, Sept. 2003, vol. 3, pp. 735–738.
- [6] M. van Herk, J. C. de Munck, J. V. Lebesque, S. Muller, and C. Rasch, “Automatic Registration of Pelvic Computed Tomography Data and Magnetic Resonance Scans Including a Full Circle Method for Quantitative Accuracy Evaluation,” *IEEE Comm. Magazine*, vol. 37, no. 12, pp. 82–88, Dec. 1999.
- [7] X. Dai and S. Khorram, “The Effects of Image Misregistration on the Accuracy of Remotely Sensed Change Detection,” *IEEE Trans. on Geoscience and Remote Sensing*, vol. 36, no. 5, pp. 1566–1577, Sept. 1998.
- [8] “When in Monterey...” Spot the Differences puzzle. [Online.] Available: <http://www.allstarpuzzles.com>
- [9] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*, Pearson Higher Education, Inc., Upper Saddle River, NJ, 2010.