

Using FPGAs to Perform Embedded Image Registration

Brandyn White
School of Electrical Engineering and Computer Science
University of Central Florida
Orlando, FL
bwhite@cs.ucf.edu

March 2008

1 Introduction

In computer vision, image registration is a technique used to find the relationship between two images of the same planar region in the world taken from different viewpoints due to the motion of a camera over time. Without a relationship of the coordinate space of one image to another, temporal information will not be fully exploitable, which generally reduces the applicable techniques to those that require only a single image to perform its function. In problems such as background subtraction, spatiotemporal information is commonly used [6, 22] to estimate the probability density function of each pixel in order to identify regions of interest in the scene, and without a pixel-to-pixel relationship between warped images this method cannot be applied. Tracking of objects found through detection is also complicated by camera motion as the motion of the object and camera are combined. At times, the spatial information provided by one image is not enough to perform the desired task, this can be remedied through the use of mosaicing, the merging of different images after they have been aligned, which allows for a moving camera to produce an increased field of view of a planar region. A similar problem is to increase the spatial resolution of an image through methods of super-resolution, which commonly use multiple images and image registration[13] to relate them to create an image of finer detail than those provided.

Currently, image registration is being applied to many diverse problems: medical imaging[3], robotics, and image stabilization[11]. It is common in robotics that the results of the registration are used in autonomous control; therefore, the results must be produceable at the same rate that they are input (i.e., real-time). As image registration is generally not real-time, it is common to make assumptions (e.g., camera motion is only translation), approximations (e.g., use registration from down-sampled images), simplifications (e.g., reduce the input image size and frame-rate), and hardware architecture optimizations (e.g., DSPs).

The goal of this research project is to develop an accurate real-time embedded image registration system that can be used in various problem domains (e.g., wearable computing, robotics, and surveillance) while investigating the natural tradeoffs between throughput, cost, generality, and accuracy. To design such a system, an image registration algorithm and complementing hardware architecture will be developed that can perform the algorithm in real-time and produce accurate results across problem domains and deliver the results in multiple formats to facilitate integration with other embedded image processing systems.

2 Embedded Image Registration

Image registration is a natural application to a specialized embedded architecture as it is often required to perform real-time, it is well understood, and is useful to include within the camera package. Simple variants appear in digital camcorders to reduce small image jitter while other devices have also been used to reduce camera shaking in surveillance applications[11], a less constrained problem. Feature point detection is the most computationally expensive operation in feature based registration, has a complex control flow,

and the algorithm choice is commonly application dependent, making it difficult to design a general purpose architecture; an affine-invariant feature detector architecture was implemented in [4]. Area based registration, with its simple control-flow, highly parallel operations, predictable memory requirements, and high accuracy make it a good candidate for embedded FPGA application. There has been work done in [10] to create an area based registration system for a UAV that uses an FPGA to create a warped image given an input image with a transformation and perform sum of squared differences (SSD) on the result to a target image, with the rest of the processing done on a desktop computer.

3 Field Programmable Gate Arrays

Originally used as a flexible method of connecting application specific integrated circuits (ASICs), field programmable gate arrays (FPGAs) have evolved considerably and now support entire designs on one chip. An FPGA is a device that has a matrix of configurable look-up tables and interconnects that allow for digital logic to be implemented with reduced cost and development time as compared to ASICs due to their programmability and provided infrastructure. A trend has been to include other specialized digital and analog devices on the chip to allow for system-on-a-chip (SoC) designs such as digital signal processing (DSP) elements, processor cores, clock managers, and memory. Modern processors commonly come with several cores which allow for different tasks to be run at the same time; in contrast, every look-up table on an FPGA can operate in parallel, thus enabling the exploitation of massive parallelism found in many hard real-world problem domains: cryptography, computer vision, bioinformatics, and speech recognition. Though they commonly have lower operating frequencies than modern computer processors, the flexible architecture provided by FPGAs allow for application specific pipelined designs which allow for increased throughput at the cost of a slight increase in latency.

3.1 FPGA Applications

The benefits of using FPGAs as a target for an application are that they are small, relatively inexpensive, provide a parallel architecture, can be reconfigured on-the-fly, and serve as an intermediate step between DSP and ASIC solutions. Of these, being able to be reconfigure itself is something that cannot be done in hardware by other devices and has been shown to be useful in several applications[5].

4 Registration Application

Image registration presents several problems when considering a hardware architecture implementation: high memory bandwidth requirements, computational expense, vulnerability to input noise, and commonly real-time operating requirements. A benefit of using FPGAs is that multiple memory devices can be used to increase bandwidth and specialized memory (e.g., ZBT, on-chip, and DDR) can be used to improve performance. Recently, DSP elements have been included on the FPGA chip to reduce the number of LUTs used to perform mathematics, these take up less power, can perform floating point math, and run close to ASIC speeds. Another addition to modern FPGAs are processors located on-chip that can run embedded operating systems, this can help considerably with input and output controls which are likely to change between applications. For example, it may be useful to output a registered mosaic to a computer over ethernet in one application and in another it may be useful to output the transformation matrices over a serial bus; using an on-chip operating system makes this modification possible using just software, avoiding complex architecture modification.

The goal of the proposed research project is to investigate the design issues and hardware application of an image registration algorithm on a FPGA for use in embedded applications. Such a device will make it possible to perform many image processing tasks in the same enclosure as the camera that were previously done off-line or through remote processing, eliminating the complexities introduced by signal transmission (e.g., noise and latency) and allowing for the information derived from the image registration results to be used in autonomous control.

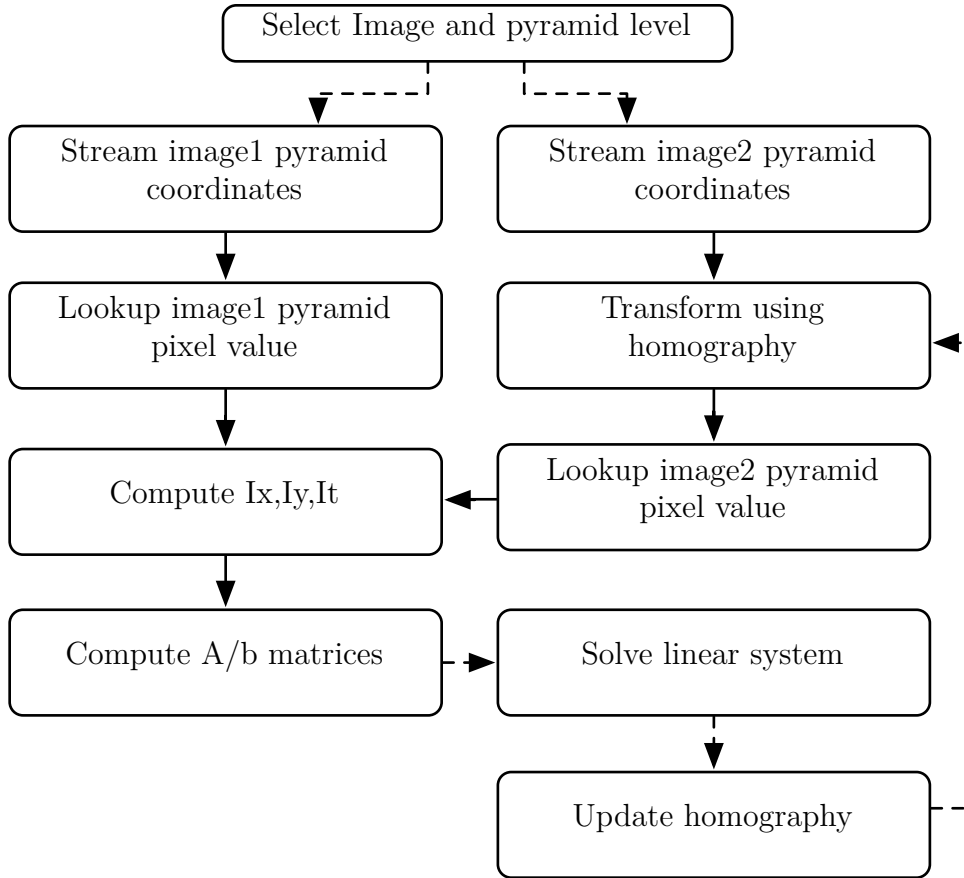


Figure 1: Architecture overview for direct affine image registration.

Two image registration algorithms are of specific interest for this task[2][16]; the former recovers only the affine (i.e., rotation, skew, and translation with 6 degrees of freedom) parameters of the homography between the images while the latter extends that method to recover the projective (i.e., affine plus the line at infinity, representing all 8 degrees of freedom) parameters of the homography. The choice of these two algorithms is motivated by the overall goal of full projective parameter estimation, while allowing for the more straightforward initial implementation of affine parameter estimation. The large overlap between these two algorithms allows for the majority of the architecture to be reused when transitioning between these two distinct phases of the project. The FPGA board to be used for this project is the Xilinx ML506 which hosts a Xilinx Virtex5-SX50 FPGA, a board designed with state-of-the-art embedded image processing applications in mind. To evaluate the performance of the architecture, publicly available data-sets will be used to compare accuracy, speed, and reliability to both of the reference implementations.

The proposed high level architecture for direct affine image registration is shown in Figure 1. The solid lines represent high bandwidth pixel data-paths and the dashed lines show low bandwidth control signals. A control structure selects the images to be registered and the pyramid level to begin the process at. The pixel coordinates for the first image are streamed and immediately read from memory; however, the second image uses an affine homography to warp the coordinates before they are read from memory. This warping is what we are attempting to iteratively estimate in this process and will align the second image with the corresponding region in the first image. From the two image value streams we compute the spatial and temporal derivatives which are then used to find the least squares solution to the affine homography minimizing the error of the brightness constancy equation. The resulting homography is then transferred to the earlier coordinate warping stage where it will be used as the process continues for several iterations and

image pyramid levels.

References

- [1] K; Hunter Appiah. A single-chip fpga implementation of real-time adaptive background model. *Field-Programmable Technology, 2005. Proceedings. 2005 IEEE International Conference on*, pages 95–102, 11.
- [2] J Bergen, P Anandan, K Hanna, and R Hingorani. Hierarchical model-based motion estimation. *Computer Vision–ECCV’92: Second European Conference on . . .*, Jan 1992.
- [3] L BROWN. A survey of image registration techniques. *portal.acm.org*.
- [4] C Cabani and W MacLean. A proposed pipelined-architecture for fpga-based affine-invariant feature detectors. *Proceedings of the 2006 Conference on Computer Vision and . . .*, Jan 2006.
- [5] R DeMara and K Zhang. Autonomous fpga fault handling through competitive runtime reconfiguration. *Evolvable Hardware*, Jan 2005.
- [6] A Elgammal, D Harwood, and L Davis. Non-parametric model for background subtraction. *FRAME-RATE Workshop*, Jan 1999.
- [7] M Fischler and R Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and *Communications of the ACM*, Jan 1981.
- [8] M Gleicher. Projective registration with difference decomposition. *IEEE Conf. of Computer Vision and Pattern Recognition*, Jan 1997.
- [9] R; Tahir Hafiz. Hardware design for registration of aerial video imagery. *Multitopic Conference, 2004. Proceedings of INMIC 2004. 8th International*, pages 682–687, 24.
- [10] Rehan Hafiz, M Ali Tahir, Omair Arshad, and Shoab A Khan. An optimized hardware accelerator for real time registration of aerial video imagery and its applications. *icig*, 00:196–199, 2004.
- [11] M Hansen, P Anandan, K Dana, G van der Wal, and et al. Real-time scene stabilization and mosaic construction. *Applications of Computer Vision*, Jan 1994.
- [12] R Hartley and A Zisserman. Multiple view geometry in computer vision. *books.google.com*, Jan 2003.
- [13] M Irani and S Peleg. Improving resolution by image registration. *CVGIP: Graphical Models and Image Processing*, Jan 1991.
- [14] C Johnston, K Gribbon, and D Bailey. Implementing image processing algorithms on fpgas. *Proceedings of the Eleventh Electronics New Zealand . . .*, Jan 2004.
- [15] D Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, Jan 2004.
- [16] S Mann and R Picard. Video orbits of the projective group a simple approach to featureless estimation of parameters. *Image Processing*, Jan 1997.
- [17] J Matas, O Chum, M Urban, and T Pajdla. Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing*, Jan 2004.
- [18] P McCurry, F Morgan, and L Kilmartin. Xilinx fpga implementation of a pixel processor for object detection applications. *Proc. Irish Signals and Systems Conf*, Jan 2000.
- [19] Mahmoud Meribout, Lazher Khriji, and Mamoru Nakanishi. A robust hardware algorithm for real-time object tracking in video sequences. *Real-Time Imaging*, 10(3):145–159, 2004.

- [20] K Mikolajczyk and C Schmid. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine ...*, Jan 2005.
- [21] Jozias Oliveira, André Printes, R C S Freire, Elmar Melcher, and Ivan S S Silva. Fpga architecture for static background subtraction in real time. pages 26–31, 2006.
- [22] C Stauffer and W Grimson. Adaptive background mixture models for real-time tracking. *Proceedings of the IEEE Computer Society Conference on ...*, Jan 1999.
- [23] J Wang, E ADELSON, V, Modeling Group, and et al. Spatio-temporal segmentation of video data. *persci.mit.edu*, Jan 1994.