MASTERAL THESIS PROPOSAL

High-performance Video Super-resolution through FPGA

Author:
Reich CANLAS

Adviser:
Engr. Carlo OCHOTORENA

A thesis submitted in fulfilment of the requirements
for the degree of Master of Science

in

Electronics and Communications Engineering

April 2015

Declaration of Authorship

I, Reich CANLAS, declare that this thesis titled, 'High-performance Video Super-resolution through FPGA' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: _____

Date: _____

"Thanks to my solid academic training, today I can write hundreds of words on virtually any topic without possessing a shred of information, which is how I got a good job in journalism."

Dave Barry

DE LA SALLE UNIVERSITY

Abstract

Gokongwei College of Engineering
Electronics and Communications Engineering

Master of Science

High-performance Video Super-resolution through FPGA

by Reich CANLAS

The Thesis Abstract is written here (and usually kept to just this page). The page is kept centered vertically so can expand into the blank space above the title too...

Acknowledgements

The acknowledgments and the people to thank go here, don't forget to include your project advisor...

Contents

List of Figures

List of Tables

Abbreviations

LAH **L**ist **A**bbreviations **H**ere

Physical Constants

Speed of Light $c = 2.997\,924\,58 \times 10^{8}\ \mathrm{ms}^{-\mathrm{S}}$ (exact)

Symbols

a	distance	m
P	power	W (Js^{-1})
ω	angular frequency	$rads^{-1}$

For/Dedicated to/To my. . .

Chapter 1

INTRODUCTION

1.1 Background of the Study

The search for ever higher screen resolutions led to the advent of ultra high definition television screens and monitors. However, to date, Common video sources do not use the full capability of most high-resolution televisions of today Super-resolution is the process of rendering or recovering a larger image or video given some low-resolution source (Dong, Loy, & He, 2014). Super-resolution finds its applications in diverse fields of study. Examples include video surveillance, in Caner, a.M. Tekalp, and Heinzelman (2003) and Zhang, Zhang, Shen, and Li (2010), medical imaging [3,4], and satellite imaging [5–7], have gained significant interests in consumer electronics, academia, and industries (temporary citation) (Cheng, Hwang, Jeng, & Lin, 2013) Image super-resolution methods that use a set of LR images construct an HR image by exploring the spatial correlations in that set (Cheng et al., 2013). Ishizaka et al. (2013) noted that "high quality SR that can be used for enterprise systems such like broadcasting and video surveillance is very compute intensive. Realtimeness is recognized as important threshold to use SR for a variety of systems. However, for example, it is several times slower than realtime to convert SD to HD by using a commodity server."

Video super-resolution is a difficult task, especially when given severe time constraints (near real-time). Today's state-of-the-art video super-resolution methods run on a GPU (high cost, high wattage). Current solutions also have unstable frame rates (Wu, Xiang, & Lu, 2011). FPGAs are known to have high performance-per-watt ratios.

1.2 Statement of the Problem

We are then confronted by the problem of finding a video super-resolution system that uses a fast algorithm to generate high-quality hi-resolution videos from a low-resolution source while maintaining a relatively small power footprint.

1.3 Objectives of the Study

1.3.1 General Objective

This study aims to come up with a new video super-resolution algorithm and implement this for use on an FPGA (field programmable gate array).

1.3.2 Specific Objectives

Specifically, the study aims on improving on the state-of-the-art video super-resolution algorithm in terms of PSNR (peak signal-to-noise ratio) and processing time Determine the FPGA best suited for the purpose of video super-resolution, considering processing resources and power consumption Profile the video super-resolution algorithm to determine performance bottlenecks Exploit parallelizable steps in the algorithm to further enhance suitability on an FPGA

1.4 Scope and Limitations of the Study

Since previous studies (cite something) have shown that clear upscaling is practical only at a factor of 4, it has been decided that this study should concern itself with x4 upscaling only.

Chapter 2

REVIEW OF RELATED LITERATURE

2.1 Image Super-resolution

Still-image super-resolution (SR) is the reconstruction of a high-resolution (HR) image given one, or a set of, low-resolution (HR) images. Super-resolution began as the problem of image restoration from a noisy signal (Helstrom, 1967). The first known work that directly tackled SR is that of Tsai and Huang (1984). Traditionally, super-resolution of images is performed with several observed LR images. This is done in order to remove artifacts introduced by the low-resolution camera sensor (Yang & Huang, 2010). There is another approach which involves only a single observation or image. The limited set of data severely limits the quality obtainable, thus the SR problem becomes ill-posed (Yang & Huang, 2010).

Super-resolution is necessary in the following fields of interest (Yang & Huang, 2010):

- Surveillance video:
- Satellite imaging:
- Medical imaging:
- Video upscaling:

2.1.1 Measuring image quality

Currently, there are two primary measures of output image quality in SR, namely, the Peak Signal-to-Noise Ratio (PSNR) and the Structural Similarity Index Measure (SSIM). The choice of PSNR or SSIM is typically arbitrary, with a few informal arguments favoring one or the other (Farsiu, Robinson, Elad, & Milanfar, 2004). To aid in selection of a suitable metric, an analysis of both image metrics is found in Horé and Ziou (2010). They state that a mathematical relationship exists between the two metrics, thus making it possible to predict the PSNR from the SSIM and vice-versa. They only differ in their sensitivity to image degradations as introduced by noise, compression, and hardware limitations.

To the present day, super-resolution remains an active area of research. The following sections present various approaches to SR that rely on several different models.

2.1.2 Image Observation Model

Several factors affect the output of a digital system, including finite aperture size and finite sensor size (Yang & Huang, 2010).

2.1.3 Frequency Domain

The first SR paper as authored by Tsai and Huang (1984) describes the SR process in the frequency domain. Their algorithm takes advantage of the shift and aliasing properties of the continuous and discrete Fourier transforms, given a set of multiple shifted low resolution images. A few extensions have been proposed, such as (Yang & Huang, 2010)

2.1.4 Interpolation-restoration

(Yang, Wright, Huang, & Ma, 2010)

2.1.5 Statistical Methods

Hello (Yang & Huang, 2010)

2.1.6 Set-theoretic Methods

Hello (Yang & Huang, 2010)

2.1.7 Dictionary Learning Methods

Digital signal data take up too much storage space, but most of this space does not account for the most significant components of the signal it represents. Compression and alternative representations are therefore required to reduce storage size while preserving fidelity. The use of orthogonal and bi-orthogonal bases in Dictionary learning is the process of training a set of mutually orthogonal basis vectors in order to create a dictionary matrix. This matrix can then model any signal as a combination of its columns, better known as ”atoms”. (citation here)

Wright, Huang, Yang, and Ma (2010) jointly trained a dictionary for low resolution and another for high resolution patches to enforce sparse representation similarity for both patch spaces. Their approach is also robust to noise, as it uses local sparse modeling. Yang, Wang, Lin, Cohen, and Huang (2012) similarly stressed the importance of learning two coupled dictionaries, (observation dictionary and latent dictionary). However, the difference in their methods is that they used a coupled dictionary learning method for single-image SR.

2.1.8 Computational Intelligence Methods

So far, the previous methods mentioned all have solid mathematical foundations. However, it has been found out (citation here) that a great number of real-world problems cannot be modeled into well-posed mathematical problems, including super-resolution. A class of algorithms dubbed ”computational intelligence” rely on mimicking natural systems to model and solve such kinds problems.

Dong et al. (2014) used a deep convolutional neural network in order to

2.2 Challenges in Image SR

Researchers still struggle with the following challenges, despite years of research.

- Image Registration
- Computation Efficiency
-

2.2.1 Image Registration

Image registration is the process of mapping two images both spatially and with respect to intensity (Brown, 1992). It is a computationally-intensive task (Yang & Huang, 2010) According to (cite here), image registration is required for the following purposes

- Integrating information taken from different sensors
- finding changes in images taken at different times or under different conditions
- Inferring three dimensional information from images in which either the camera or the objects in the scene have moved
- Model-based object recognition

2.3 Video super-resolution

Video super-resolution is the application of super-resolution to moving pictures. It can generally be divided into two categories: incremental and simultaneous (Su, Wu, & Zhou, 2011). The former category is faster but less visually consistent to the human eye. Liu and Sun (2014) mentions that video SR is relatively more challenging than image SR which has been studied for decades. They also propose a Bayesian video SR system that can simultaneously estimate the motion, blur kernel, and noise level.

2.3.1 Mathematical Methods

2.3.2 Computational Intelligence Methods

2.4 High Performance Computing Platforms

Yang and Huang (2010) suggest that high-performance hardware does matter in tackling super-resolution problems. In (insert citation here), something. Modern CPUs (central processing units) of computers combine high-frequency processors with a degree of parallelism to add more processing power to algorithms. Even so, the CPU is not enough to handle tasks such as SR in real-time. There are several steps in the SR process that may be implemented as parallel tasks. Following are the discussions on GPUs, manycore coprocessors, and FPGAs, three parallel platforms commonly in use today.

2.4.1 Graphics Processing Units

General-purpose computing on Graphics Processing Units GPUs (Graphics Processing Units) have been favored in recent years for this task, as it offers high amounts of parallelism (due to its multiple cores) and compatibility with existing computer systems and programming paradigms. Wu et al. (2011) claims 6x speedup against the same algorithm implemented on a CPU. Shen, Wu, and Deng (2014) used a real-time learning-based SR algorithm based on error feedback.

2.4.2 Manycore Coprocessors

This class of parallel processors are based off CPU architectures but have more cores than the traditional CPU and are meant to run at a lower frequency. A host CPU passes the appropriate parallel instructions to the manycore coprocessor and subsequently fetches the results of the computation. Manycore processors offer more programmability than GPUs simply by the fact that they share the same architecture as the host CPU (Ishizaka et al., 2013).

2.4.3 Field Programmable Gate Arrays (FPGAs)

FPGAs (Field Programmable Gate Arrays) are logic devices that can be reconfigured by a designer on the field after being manufactured. Since at the lowest level, logic circuits are inherently parallel and real-time, FPGAs offer optimization potential that cannot be realized when using instruction-based platforms such as CPUs and GPUs. FPGAs typically run at much lower frequencies than CPUs and GPUs, making them more power efficient. Higher-end FPGAs even offer the ability to be partially dynamically reconfigured, so that even while it is running, parts of the FPGA fabric gets their design altered (Dye, 2012). These factors makes FPGAs suitable for the most computationally-intensive real-time applications.

2.4.3.1 Design Strategies

Sirowy and Forin (2008) detailed the reasons why an FPGA offers high speedups over instruction-based processors such as CPU, manycore, and GPU. Among these are the removal of an instruction fetch step,

2.4.3.2 Use in super-resolution applications

Angelopoulou, Bouganis, Cheung, and Constantinides (2009) created a real-time video SR system on an FPGA that is robust against noise. It uses the iterative back projection algorithm.

2.4.3.3 As video processors

Roth (2011) used low-cost FPGA hardware to accomplish real-time video processing tasks such as deinterlacing, alpha blending, and frame buffering.

Chapter 3

CONCEPTUAL FRAMEWORK

3.1 Research Process

3.1.1 Evaluation of state-of-the-art algorithms

3.1.2 Modifications for speed and quality

Morbi rutrum odio eget arcu adipiscing sodales. Aenean et purus a est pulvinar pellentesque. Cras in elit neque, quis varius elit. Phasellus fringilla, nibh eu tempus venenatis, dolor elit posuere quam, quis adipiscing urna leo nec orci. Sed nec nulla auctor odio aliquet consequat. Ut nec nulla in ante ullamcorper aliquam at sed dolor. Phasellus fermentum magna in augue gravida cursus. Cras sed pretium lorem. Pellentesque eget ornare odio. Proin accumsan, massa viverra cursus pharetra, ipsum nisi lobortis velit, a malesuada dolor lorem eu neque.

3.1.3 FPGA Deployment and Comparative Tests

3.1.4 Exploiting parallelism for FPGAs

3.2 Data Flow

Chapter 4

METHODOLOGY

Chapter 5

SUMMARY

Appendix A

Gantt Chart

Write your Appendix content here.

Appendix B

Proposed Budget

Write your Appendix content here.

References

Angelopoulou, M. E., Bouganis, C.-S., Cheung, P. Y. K., & Constantinides, G. a. (2009). Robust Real-Time Super-Resolution on FPGA and an Application to Video Enhancement. *ACM Transactions on Reconfigurable Technology and Systems*, *2*(4), 1–29. doi: 10.1145/1575779.1575782

Brown, L. G. (1992). A survey of image registration techniques. *ACM Computing Surveys*, *24*(4), 325–376. doi: 10.1145/146370.146374

Caner, G., a.M. Tekalp, & Heinzelman, W. (2003). Super resolution recovery for multi-camera surveillance imaging. *2003 International Conference on Multimedia and Expo. ICME '03. Proceedings (Cat. No.03TH8698)*, *1*. doi: 10.1109/ICME.2003.1220866

Cheng, M. H., Hwang, K. S., Jeng, J. H., & Lin, N. W. (2013). Classification-based video super-resolution using artificial neural networks. *Signal Processing*, *93*(9), 2612–2625. Retrieved from http://dx.doi.org/10.1016/j.sigpro.2013.02.013 doi: 10.1016/j.sigpro.2013.02.013

Dong, C., Loy, C. C., & He, K. (2014). Image Super-Resolution Using Deep Convolutional Networks. *arXiv preprint*, 1–14.

Dye, D. (2012). *Partial Reconfiguration of Xilinx FPGAs Using ISE Design Suite* (Vol. 374).

Farsiu, S., Robinson, D., Elad, M., & Milanfar, P. (2004). Advances and challenges in super-resolution. *International Journal of Imaging Systems and Technology*, *14*(2), 47–57. doi: 10.1002/ima.20007

Helstrom, C. W. (1967). Image Restoration by the Method of Least Squares. *Journal of the Optical Society of America*, *57*(3), 297. doi: 10.1364/JOSA.57.000297

Horé, A., & Ziou, D. (2010). Image quality metrics: PSNR vs. SSIM. *Proceedings - International Conference on Pattern Recognition*, 2366–2369. doi: 10.1109/ICPR.2010.579

Ishizaka, K., Miyamoto, T., Akimoto, S., Iketani, a., Hosomi, T., & Sakai, J. (2013). Power efficient realtime super resolution by virtual pipeline technique on a server

with manycore coprocessors. *IEEE Symposium on Low-Power and High-Speed Chips - Proceedings for 2013 COOL Chips XVI*, *2*, 2–4. doi: 10.1109/CoolChips .2013.6547918

Liu, C., & Sun, D. (2014). On bayesian adaptive video super resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *36*(2), 346–360. doi: 10.1109/TPAMI.2013.127

Roth, F. (2011). *Using low cost FPGAs for realtime video processing* (Unpublished doctoral dissertation). MASARYK UNIVERSITY.

Shen, Y., Wu, X., & Deng, X. (2014, December). GPU-aided real-time image/video super resolution based on error feedback. In *2014 ieee visual communications and image processing conference* (pp. 286–290). IEEE. Retrieved from http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7051560 doi: 10.1109/VCIP.2014 .7051560

Sirowy, S., & Forin, A. (2008). Where's the Beef ? Why FPGAs Are So Fast. *Microsoft Research*(September).

Su, H., Wu, Y., & Zhou, J. (2011). Adaptive incremental video super-resolution with temporal consistency. *Proceedings - International Conference on Image Processing, ICIP*, 1149–1152. doi: 10.1109/ICIP.2011.6115632

Tsai, R., & Huang, T. S. (1984). Multiframe image restoration and registration. *Advances in computer vision and Image Processing*, *1*(2), 317–339.

Wright, J., Huang, T., Yang, J., & Ma, Y. (2010). Image Super-Resolution via Sparse Representation. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, 1–13. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20483687$\delimiter"026E30F$nhttp://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4587647 doi: 10.1109/TIP.2010.2050625

Wu, X., Xiang, H., & Lu, P. (2011). A GPU Accelerated Algorithm for Compressive Sensing Based Image Super-Resolution. *2011 Workshop on Digital Media and Digital Content Management*(60633030), 198–202. doi: 10.1109/DMDCM.2011 .10

Yang, J., & Huang, T. (2010). Image super-resolution: Historical overview and future challenges. In *Super-resolution imaging* (pp. 3–25). Retrieved from http://books.google.com/books?hl=en&lr=&id=fjTUbMnvOkgC&oi=fnd&pg=PA1&dq=Image+super-resolution:

+Historical+overview+and+future+challenges&ots=53GZConOGy&sig=oHgGJVo_57Tv19cxLV_XMtC1OPw

Yang, J., Wang, Z., Lin, Z., Cohen, S., & Huang, T. (2012). Coupled dictionary training for image super-resolution. *IEEE Transactions on Image Processing*, *21*(8), 3467–3478. doi: 10.1109/TIP.2012.2192127

Yang, J., Wright, J., Huang, T., & Ma, Y. (2010). Image Super-Resolution Via Sparse Representation. *IEEE Transactions on Image Processing*, *19*(11), 2861–2873. Retrieved from http://ieeexplore.ieee.org/ielx5/83/5602158/05466111.pdf?tp=&arnumber=5466111&isnumber=5602158$\delimiter"026E30F$nhttp://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5466111 doi: 10.1109/TIP.2010.2050625

Zhang, L., Zhang, H., Shen, H., & Li, P. (2010). A super-resolution reconstruction algorithm for surveillance images. *Signal Processing*, *90*(3), 848–859. Retrieved from http://dx.doi.org/10.1016/j.sigpro.2009.09.002 doi: 10.1016/j.sigpro.2009.09.002