

95pJ/Label FPGA 低功耗光場廣域深度處理器

A 95pJ/label Wide-Range Depth-Estimation Processor for Full-HD Light-Field Applications on FPGA

DESIGN GROUP D19-085

隊伍名稱 就剩我了
The Last of Us

隊長 陳立得 / 清華大學電機工程研究所



指導教授

黃朝宗
清華大學電機工程學系

臺灣大學電子工程學研究所博士，現為清華大學電機工程學系副教授。曾服務於聯詠科技，亦曾於麻省理工學院與臺灣大學進行博士後研究。

研究領域

近來研究以實現高效能、高品質之電腦視覺與計算攝影學應用為主，研究內容橫跨演算法與晶片設計，是極少數能同時發表頂尖會議論文於電腦視覺領域 (CVPR、ICCV) 以及固態電子電路領域 (ISSCC、VLSI Symp.) 之學者。

高解析度與廣深度範圍的深度圖是光場渲染及機器視覺領域諸如視角合成、背景虛化或是三維場景重建、擴增實境等應用不可或缺的核心資訊。隨著創新應用出現以及對於拍攝影像品質的追求，近年來移動端裝置手機平板上也逐漸出現了可以用於深度估計的多鏡頭相機系統。利用立體視覺的原理，多鏡頭相機系統，可以利用不同視角的影像推算出場景的深度圖。然而傳統的深度估計演算法在應用於多鏡頭系統時，會有運算量過高、DRAM 存取過於頻繁、功耗過高等限制，使得在行動裝置上難以推廣。

為了探索多鏡頭影像系統之深度估計，並使得這樣的系統能夠有效率地裝載在移動裝置上，我們提出了改進的深度估計演算法，並針對所提出的演算法設計了對應的硬體架構，以 FPGA 完成一個低功耗的光場深度處理器。配合所提出的演算法及低功耗設計技巧，我們能夠在 Xilinx ZC706 平台上完成與 ASIC 相似功耗的設計。同時我們實作了展示系統，直接透過影像輸出呈現所估計的深度圖。證明透過我們的設計，可以將複雜的五視角深度估計的硬體裝載在 FPGA SoC 平台上，並透過實際展示證明系統的可行性。

設計特點

應用場景：我們的設計能夠應用於各種多鏡頭的相機系統中，包含 2、3、5 鏡頭的相機系統，可以在不同的應用場景，平台功耗限制下提供合適的深度估計結果。並且支援廣範圍的深度搜尋，適用於高解析度的影像深度估計。而所提出的演算法硬體架構，更能減少 DRAM 存取需求與功耗，適用於移動裝置平台。

規格

- 提供每秒 30 張，128 Pixel 範圍 (6.6% Full HD) 的 Full HD 解析度深度圖
- 功耗：384 mW

演算法與硬體架構共同設計

- Octave Depth Sampling：重新設計深度搜尋的標籤，提高深度搜尋範圍，並使得光場影像能以影像金字塔的形式存取
 - 擴展深度搜尋範圍達到 128 Pixel

- 節省 53% DRAM 存取頻寬
- 節省 31% 的系統內記憶體使用
- Image-Guided Inference 整合深度估計最佳化及後處理流程，使得低解析度下的深度估計結果，能夠在影像作為引導下，得到高解析度、高品質的深度圖
 - 降低了 75% 的運算複雜度

實作成果

我們的設計在 Xilinx ZC706 上實作佔用 167K 的 LUT 以及 329K bytes 的 BRAM 記憶體資源。我們的系統操作在 54 MHz，可以在 384 mW 的功率消耗下提供每秒 30 張 Full HD 五視角的深度圖估計。平均每個深度標僅需 95 pJ，在 FPGA 上提供與 ASIC 相似的功率消耗效率。



圖 1. 五鏡頭相機系統示意圖，及深度估計處理器在 FPGA SoC 平台上運作展示

ADDITIONAL

High-resolution and wide-range depth maps are the key to enable novel light-field and computer vision applications, such as digital refocusing, and view synthesis, or 3D reconstruction, and augmented reality. In recent years with the demand for high-quality image/video acquisition and novel applications, mobile platforms have moved to a dual-camera or even multi-camera system. Based on stereo matching, a multi-camera system can estimate the depth information from the captured scenes. Unfortunately, applying traditional stereo matching algorithms on multi-camera systems requires complex computations, high DRAM accessing bandwidth, and will result in high power consumption.

To explore the possibility of the depth estimation for multi-camera systems, and efficiently deploy such complex depth processor on mobile platforms, we present an energy-efficient depth estimation algorithm with corresponding hardware architecture to meet this purpose. The design is implemented on Xilinx ZC706 Soc Platform and achieve ASIC-comparable energy efficiency-95pJ/label- for Full-HD five-view light fields (multi-camera) at 30fps.

Design Features

Application:
Our design is compatible with various multi-camera systems, including 2, 3, 5 cameras; and can provide proper depth estimation result based on different application and platform constraints.

Our design supports a wide depth search range, which makes it suitable for high-resolution applications. Furthermore, the proposed algorithm-hardware codesigned architecture reduces the requirement to power and DRAM access, allow deploying such intricate design on mobile platforms.

Specification

- 30 fps Full HD depth estimation
- Support 128 pixels depth search range (6.6% Full HD size)
- 384 mW power consumption

Algorithm-Hardware Codesign

- Octave Depth Sampling:
We redesign the labels for depth estimation to enlarge the depth search range and allow the light field images to be stored and accessed in an image pyramid.
 - Extend the search range to 128 pixels
 - Reduce 53% DRAM accessing bandwidth
 - Reduce 31% on-Chip Memory usage
- Image-Guide Inference
Integrate the depth estimation optimization and post-processing. This technique allows us to estimate depth on a low-resolution light field while preserving the depth map quality.
 - Reduce 75% computing resource

Implementation Result

We implement the design on Xilinx ZC706 platform. It occupies 167K LUTs, 67K Registers, and 329K bytes of BRAM. Our design operates at 54 MHz and consumes 384mW. It equivalently achieves 95pJ per depth label for Full-HD 30fps.

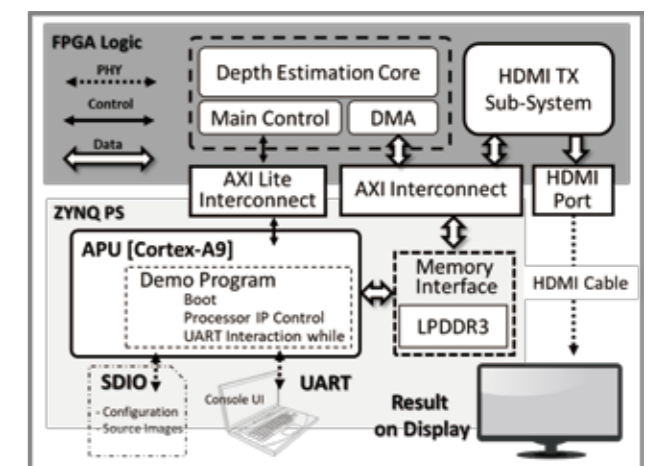


Fig.2 The System Architecture of the Domo Platform