Streszczenie

Systemy nadzoru wizyjnego stały się integralną częścią przestrzeni miejskiej, w której pracują i przebywają ludzie. Rozwój technik przetwarzania obrazów i postęp techniczny w dziedzinie kamer oraz mocy obliczeniowej procesorów spowodował, że sieci nadzoru wizyjnego są coraz większe. Klasyczne podejście, w którym dane wizyjne wysyłane są do centralnego serwera i tam analizowane, nie skaluje się do wymagań oraz rozmiaru obecnych aplikacji. Naturalnym krokiem jest przejście w kierunku systemów rozproszonych, gdzie obliczenia realizowane są przez węzły sieci nadzoru wizyjnego. Węzły, które dokonują akwizycji danych obrazowych, ich analizy, a następnie przesłania metadanych o obserwowanej scenie nazywamy kamerami inteligentnymi.

W niniejszej pracy doktorskiej zaproponowano ogólną architekturę logiczną i sprzętową kamery inteligentnej, która umożliwia realizację pogłębionego przetwarzania obrazu. Dedykowana jest dla heterogenicznych układów reprogramowalnych, które w swojej strukturze oferują logikę programowalną oraz osadzone, wysokowydajne rdzenie ogólnego przeznaczenia. Opisane rozwiązanie udostępnia moduły programowe i sprzętowe umożliwiające łatwe łączenie różnych akceleratorów sprzętowych, przeznaczonych dla szerokiego spektrum aplikacji, jednak w pracy przedstawiono rozwiązania związane z przetwarzaniem obrazów.

W ramach pracy zaimplementowano akceleratory sprzętowo-programowe realizujące kolejne etapy przetwarzania obrazów w typowym potoku przetwarzania kamery inteligentnej. Pierwszym krokiem jest detekcja ruchomych obiektów w sekwencji wideo. Wybrany algorytm przybliżonej mediany w połączeniu z dodatkowymi etapami filtracji i obróbki danych oferuje dobrą jakość, a przy tym jest efektywny w translacji do sprzętowej wersji. Następnym etapem jest detekcja, deskrypcja i dopasowanie cech punktowych. Do realizacji tego zadania zaimplementowano sprzętowe akceleratory dla detektora cech punktowych FAST, binarnego deskryptora BRIEF oraz modułu dopasowania. Rozwiązanie to w momencie publikacji było najbardziej wydajną i sprzętowo efektywną implementacją opisanych algorytmów, a dodatkowo charakteryzowało się możliwością skalowania do potrzeb aplikacji poprzez elastyczny moduł dopasowania. Ostatnim etapem przetwarzania jest klasyfikacja obiektów. Wykorzystano deskryptor obszarów LBP w połączeniu z klasyfikatorem bazującym na drzewach decyzyjnych. Przy implementacji przedstawionych akceleratorów wykorzystano zaproponowaną uniwersalną architekturę.

Poza rozwiązaniami sprzętowymi w pracy przedstawiono rozwiązania programowe, jak nową metodę sterowania przepływem informacji oraz aktywnością w sieci kamer inteligentnych. Bazuje ona na wymianie informacji o aktywności na obserwowanych scenach w celu obniżenia zużycia energii całej sieci. Dodatkowo w celu ułatwienia wykorzystania zaproponowanej architektury sprzętowej opracowano warstwową architekturę programową, która ułatwia wdrażanie i testowanie nowych algorytmów w sieciach sensorów wizyjnych. Zaproponowany klasyfikator oparty o drzewa decyzyjne został wygenerowany automatycznie przy użyciu dedykowanego oprogramowania, który pozwala w łatwy sposób opracować implementację sprzętową, dedykowaną dla układów FPGA, drzewa decyzyjnego lub lasu losowego dla dowolnego problemu. Wszystkie opracowane rozwiązanie programowe zostały upublicznione w postaci wolnego oprogramowania.

W ramach prac nad algorytmami przetwarzania obrazu wykorzystywanymi w kamerze inteligentnej i systemach nadzoru wizyjnego zebrano zestaw danych testowych, które zostały oznaczone i udostępnione jako otwarta baza danych.

Zintegrowanie rozwiązań opracowanych w ramach pracy doktorskiej w heterogenicznym układzie reprogramowalnym pozwala na uzyskanie kamery inteligentnej o wysokich parametrach użytecznych. Dzięki realizacji większości, obliczeniowo kosztownych, zadań jako akceleratorów sprzętowych zaimplementowanych w logice programowalnej, zużycie energii jest na niskim poziomie zważywszy, na to jak zaawansowane przetwarzanie jest realizowane. Wykorzystanie zaproponowanej architektury i dostępnych modułów sprzętowych oraz programowych minimalizuje wymagany nakład pracy, co jest największym problem związanym z wykorzystaniem rozwiązań sprzętowych zamiast procesorów ogólnego przeznaczenia.

Michał Fu

Abstract

Video surveillance systems have become an integral part of the urban space where people work and live. Developments in image processing techniques, technical progress in the field of imaging and availability of more powerful processors caused the said systems to be more complex than ever before. The classic approach, in which video data is sent to a central server and analyzed there, does not scale up to the requirements and size of current applications. The natural step is the transition towards fully distributed systems, where calculations are performed by nodes of the video surveillance network. Nodes that can acquire image data, analyze it, and then send only the metadata about the observed scene are called smart cameras.

This doctoral dissertation proposes a general architecture for a smart camera, both for software and hardware parts, that enables the implementation of an in-depth image processing routines. It is dedicated to heterogeneous, reprogrammable systems, that contain both programmable logic resources and embedded high performance, general-purpose cores in them. The described solution provides software and hardware modules enabling easy connection of various accelerators, designed for a wide spectrum of applications. However, the thesis is focused on solutions designed for image processing.

As part of the work, hardware and software accelerators have been implemented to carry out subsequent stages of the typical image processing pipeline used in a smart camera. The first step is the detection of moving objects in a video sequence. The selected algorithm of the approximate median, when combined with additional filtering and some data processing stages, is robust and offers good quality. In addition to that, it easily and effectively translates into a hardware implementation. The next step is the detection, description, and matching of point features. To accomplish this task, hardware accelerators have been implemented for the FAST point detector, the BRIEF binary descriptor, and the matching module. At the time of publication, this solution was the most efficient in terms of performance and resources-usage implementation of the described algorithms in an FPGA device. Another advantage of the proposed solution is its ability to scale to the needs of the application thanks to a flexible matching module. The final stage of processing is object classification. The LBP region descriptor was used in combination with a decision tree based classifier. The proposed universal architecture of a smart camera was used to implement the presented accelerators.

In addition to hardware solutions, the thesis presents a set of software solutions, such as a new method of controlling information flow and activity in the network of smart cameras. It reduces the energy consumption of the whole network based on the information about activity on each of the observed scene send by the cameras. To ease the process of implementation of the proposed hardware architecture, a layered, software architecture was developed. It helps with testing new algorithms in video sensor networks. The proposed, decision tree based classifier has been generated automatically using dedicated software. The prepared software allows the translation of any decision tree or random forest classifier to a hardware implementation in an FPGA device. All the developed software solutions have been made public in the form of open-source software.

As part of the work on image processing algorithms designed for smart cameras and video surveillance systems, a set of test data was collected and annotated. It was then described and made available as an open database.

The integration of all the solutions presented in the thesis as a system on a chip in a heterogenous reprogrammable device gives a very capable and unique smart camera. Thanks to the implementation of the majority of the compute-intensive tasks in hardware, energy consumption is relatively small, especially when considering how advanced processing is realized. The use of the proposed architecture and available hardware and software modules largely reduces the amount of work required to move from software base solution to one that utilizes the hardware.

Michał Fule