

Artur ZAWADZKI, Marek GORGON

AGH AKADEMIA GÓRNICZO-HUTNICZA, WYDZIAŁ EAIIE, KATEDRA AUTOMATYKI,
Al. Mickiewicza 30, 30-059 Kraków

Nowa architektura wizyjnego systemu wbudowanego dedykowana dla kamery inteligentnej

Mgr inż. Artur ZAWADZKI

Absolwent kierunku Automatyka i Robotyka na Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie (2006). Ukończył studia doktoranckie w dziedzinie Automatyka i Robotyka, aktualnie przygotowuje rozprawę doktorską o charakterze aplikacyjnym dotyczącą zastosowania układu FPGA do śledzenia obiektów ruchomych. Równoległe, od 2005 roku pracuje w firmie ES-SYSTEM jako konstruktor mechatronik.



e-mail: zawadzki@agh.edu.pl

Dr hab. inż. Marek GORGON

Autor jest absolwentem Akademii Górniczo-Hutniczej w Krakowie, gdzie uzyskał kolejno dyplom magistra inżyniera elektroniki (1988) oraz stopnie doktora (1995) i doktora habilitowanego (2007), oba w dyscyplinie Automatyka i Robotyka. Pracuje w Katedrze Automatyki AGH i zajmuje się metodami i systemami do przetwarzania i analizy obrazu. Od 1992 roku jego głównym obszarem zainteresowania są systemy obrazowe oparte na układach FPGA. Jest członkiem IEEE i autorem 50 publikacji.



e-mail: mago@agh.edu.pl

Streszczenie

W artykule przedstawiono przegląd wybranych metod podłączenia kamer z interfejsem cyfrowym do systemów mikroprocesorowych lub rekonfigurowalnych, z podziałem na systemy z komputerem nadrzędnym oraz na tzw. *smart camera*. W drugim przypadku konieczne jest zadbanie o odpowiednią architekturę odpowiedzialną za przesłanie zarejestrowanego obrazu do jednostek obliczeniowych. Zaprezentowano rozwijany system oparty o układ FPGA, o nowatorskiej architekturze, zawierający dwa sensory CMOS i rozbudowany układ pamięci zewnętrznych, mogący służyć do realizacji specyficznych zadań przetwarzania obrazów.

Słowa kluczowe: przetwarzanie obrazu, FPGA, kamera inteligentna, system wbudowany.

A new architecture of the FPGA-based embedded vision system dedicated for smart cameras

Abstract

Embedded vision systems are becoming more popular. A smart camera consists of an image sensor and a computing unit processing the image. Integration of many new functions in vision systems is now possible due to progress of resources of FPGA devices. A new vision system device presented in this paper is unique with regard to its architecture. The main goal during development of this smart camera was to provide it with a possibly large number of memory banks for storing image data, while keeping compact dimensions and low price. In the presented system there was chosen processing of monochromatic images, so significant reduction in the width of memory data buses was obtained. Thanks to that, up to eight fully independent memory banks was connected to the FPGA device. The second assumption was usage of static memories, which decided in favour of large functionality of the device, because very fast data transfer from random address location was then possible. This meets requirements of image processing algorithms, which is computing data in Region of Interest, for example.

Keywords: image processing, FPGA, smart camera, embedded system.

1. Wstęp

Wbudowane systemy wizyjne stają się coraz bardziej popularne. Realizacja zadań przetwarzania i analizy obrazu w układach FPGA jest dobrze opanowana i rośnie jej popularność. Między innymi, układy FPGA wykorzystuje się w kamerach inteligentnych (ang. *smart camera*). Kamera inteligentna składa się z czujnika obrazu oraz układu przetwarzającego obraz.

Nowa koncepcja systemu wizyjnego, opartego na układzie FPGA, przedstawiona w niniejszej publikacji, jest unikalna, ze względu na swoją architekturę. Podstawowym założeniem uczynionym przy konstruowaniu kamery było zapewnienie możliwie dużej liczby banków pamięci służących do przechowywania ramek obrazu, przy zachowaniu kompaktowej budowy i niskich

kosztów wytworzenia. Mała liczba banków pamięci znacznie ogranicza możliwości implementacji złożonych operacji analizy obrazu, w których wymagany jest równoległy dostęp do kilku ramek obrazu. Rozwiązania komercyjne, oferujące wirtualne banki pamięci w układach pamięci dynamicznej o dużej pojemności, ograniczają szybki i równoległy dostęp do wskazanych lokacji adresowych. W przedstawionej architekturze, w przeciwieństwie do trendów komercyjnych, zdecydowano się na dołączenie do układu FPGA aż ośmiu niezależnie obsługiwanych banków pamięci.

Drugim ważnym założeniem, stanowiącym o oryginalności rozwiązania, jest użycie szybkich pamięci statycznych, które pozwalają na bardzo efektywne transmitowanie do układu FPGA danych o dowolnych lokacjach adresowych. Stosowanie pamięci statycznych było dość powszechne w kartach FPGA starszej generacji, jest jednak stosunkowo rzadkie w kartach akceleratorów wykorzystujących układy FPGA nowszych serii. Karta kamery inteligentnej nie jest jednak akceleratorem ogólnego przeznaczenia, lecz układem wbudowanym o dedykowanej funkcji. Dotychczasowe doświadczenia implementacji algorytmów analizy obrazów, w szczególności algorytmów różnicowych, generacji tła i podobnych, uzasadniają taką konstrukcję karty. Do innych cech urządzenia należy możliwość podłączenia dwóch elementów akwizycji obrazu. Jest to zbieżne z narastającym zainteresowaniem realizacją algorytmów stereowizyjnych.

2. Rodzaje architektur do przetwarzania i analizy obrazu w trybie ciągłym

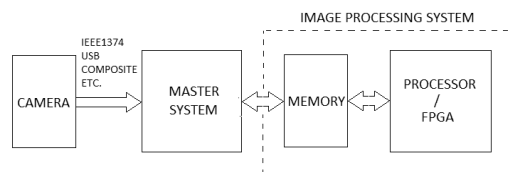
Wśród systemów przetwarzających i analizujących obraz sporą grupę stanowią systemy, które pozyskują informację wizyjną w sposób ciągły za pomocą kamer podłączonych do jednostki obliczeniowej. Sytuacja taka zachodzi najczęściej wówczas, gdy system ma analizować dane wizyjne w czasie rzeczywistym [1]. W takim przypadku konieczne jest zastosowanie konfiguracji, która zapewni uzyskanie odpowiedniej wydajności w trakcie rejestracji i przesłania obrazu do systemu obliczeniowego.

2.1. Architektura z komputerem nadrzędnym

Najprostszą metodą dającą dostęp do danych obrazowych dla systemu obliczeniowego jest bezpośrednie przesłanie bloku danych z systemu nadrzędnego wyposażonego w kamerę do podsystemu obliczeniowego (rys. 1). Jednostką nadrzędną może być komputer PC albo specjalizowane urządzenie, do którego podłączona jest kamera poprzez jeden z wielu powszechnie stosowanych standardów transmisji sygnału wizyjnego [2].

Opisana grupa rozwiązań posiada jedną zasadniczą zaletę – nie trzeba budować od podstaw podsystemu akwizycji obrazu, lecz można wykorzystać znane i powszechnie dostępne elementy oraz oprogramowanie. Przedstawiona architektura jest zalecana, jeśli

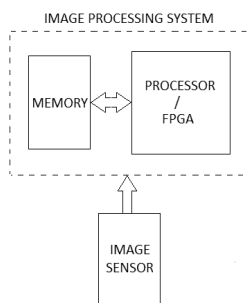
pozwała wykonać zadania obliczeniowe w czasie wystarczającym dla realizowanej aplikacji.



Rys. 1. Klasyczny system przetwarzania obrazów
Fig. 1. Scheme of classical image processing

2.2. Architektura specjalistyczna – kamera inteligentna (*smart camera*)

W zastosowaniach wymagających miniaturyzacji i redukcji złożoności konstrukcyjnej całości systemu przetwarzania i analizy obrazów, a także, gdy platforma obliczeniowa pozbawiona jest popularnych interfejsów multimedialnych lub wydajnego łącza komunikacyjnego, wcześniej opisane rozwiązanie nie jest możliwe do zastosowania. Jednym z możliwych rozwiązań jest zintegrowanie elementu, bądź elementów do akwizycji obrazu z wyjściem cyfrowym z systemem obliczeniowym (rys. 2). Unika się dzięki temu członów pośredniczących i możliwe jest zintegrowanie całego systemu w stosunkowo niewielkiej przestrzeni. W jednej obudowie mieści się wówczas kamera, pamięć, element przetwarzający (np. procesor lub układ rekonfigurowalny), układy peryferyjne i zasilające. Rozwiązania tego typu najczęściej zwane są kamerami inteligentnymi [3].



Rys. 2. Struktura kamery inteligentnej
Fig. 2. Block diagram of a smart camera

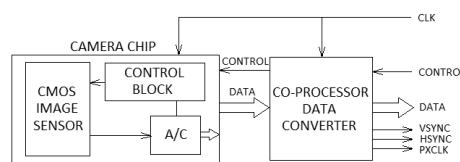
Szczególnie korzystne jest przeprowadzenie w kamerze operacji analizy obrazu. Zmniejsza to radykalnie ilość danych przesyłanych pomiędzy kamerą a systemem nadrzędnym. Do systemu nadrzędnego przesyłane są bądź przetworzone i wyselekcjonowane dane pikselowe, bądź informacja pozyskana z obrazu (np. deskryptory, lokalizacja obiektów, trajektoria ruchu).

3. Opis rozwiązań wbudowanych (*embedded*)

Kamerę z systemem wyposażonym w układ rekonfigurowalny lub mikroprocesorowy można połączyć z systemem nadrzędnym na różne sposoby, zależnie od typu kamery. Przetworniki wizyjne (ang. *vision sensors*) można podzielić na dwie główne grupy – CCD (ang. *Charge Coupled Device*) oraz CMOS (ang. *Complementary Metal Oxide Semiconductor*). Porównanie cech tych czujników przedstawiono w pracy [4].

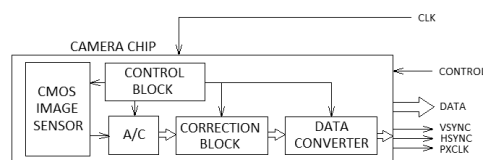
Ponieważ przetworniki CMOS posiadają wyjście cyfrowe, są dobrymi kandydatami jako sensory dla kamery inteligentnej. Najprostsze z nich zawierają oprócz sensora jedynie prosty interfejs szeregowo-równoległy i generator sygnałów sterujących, zaś zamiana na formę równoległą danych wraz z korektą wzmacnień w torze wideo odbywa się w zewnętrznym koprocesorze [5]

(rys. 3). Ten typ podłączenia jest czasem spotykany w podręcznym sprzęcie powszechnego użytku wyposażonym w kamery CMOS (np. telefony komórkowe).



Rys. 3. Podłączenie kamery z wykorzystaniem koprocesora
Fig. 3. Interconnection with co-processor

Układy o większym stopniu integracji, najbardziej popularne, zawierają w jednej strukturze krzemowej sensor, generatory przebiegów taktujących, przetworniki A/C i kontroler odpowiedzialny za sterowanie wewnętrznymi blokami. Tego typu układy wyposażone są w prosty interfejs równoległy [6] (rys. 4). Oprócz nich, w interfejsie obecne są linie synchronizacji poziomej HSYNC i pionowej VSYNC, a także sygnał synchronizujący PXCLK.



Rys. 4. Kamera z wyjściem cyfrowym
Fig. 4. Camera with digital output

W przypadku podłączenia kamery z interfejsem równoległym do systemu mikroprocesorowego, można zastosować kilka rozwiązań, zależnych od stosowanego procesora. Niektóre mikrokontrolery, np. z rdzeniem ARM9 firmy Atmel [7], posiadają interfejs zwany ISI. Interfejs ten wyposażony jest w te same linie komunikacyjne co kamera. Z pomocą wbudowanego kontrolera DMA możliwy jest transfer kompletnych ramek do pamięci wewnętrznych, na ekran LCD, bez udziału rdzenia procesora. Należy uznać, że ten rodzaj podłączenia jest zdecydowanie najłatwiejszy. Gdy mikrokontroler nie posiada interfejsu ISI, a wewnętrzne zasoby pamięciowe są wystarczające, można pokusić się o bezpośrednie podłączenie kamery do linii I/O [8]. Wymagana jest wtedy ścisła synchronizacja, co zazwyczaj oznacza konieczność oprogramowania kluczowych procedur w asemblerze. Ponadto ze względu na wymóg stosowania mniej lub bardziej zaawansowanego przetwarzania wstępnego, realizowanego równocześnie z odczytem danych, rozdzielczość przetwarzanych obrazów nie może być zbyt duża.

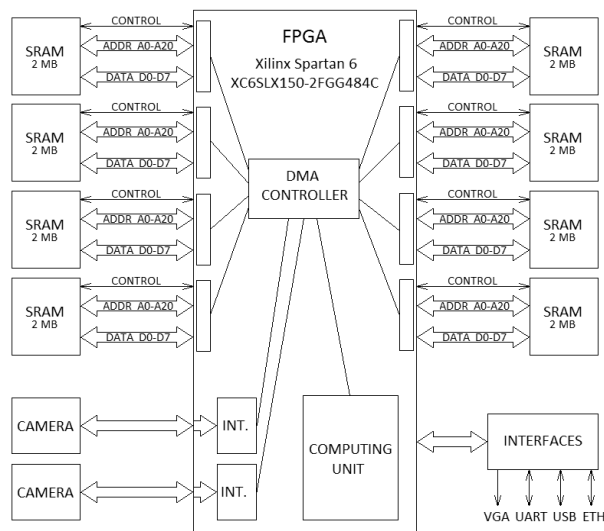
Trzecią metodą jest wykorzystanie prostej i taniej logiki pośredniczącej, służącej jako sprzęg z magistralą danych i adresową systemu mikroprocesorowego [9]. Może to być mały układ CPLD albo logika TTL/CMOS w postaci liczników i buforów. Rozwiązanie takie zapewnia, przy minimalnym udziale głównego procesora, zapis danych obrazowych w pamięci zewnętrznej (np. typu SRAM), inicjowany przez mechanizmy synchronizujące kamery. Umożliwia łatwą rozbudowę systemu o dodatkowe kamery, z niezależnym kierowaniem strumieni danych do wybranych pamięci. Wymaga jednak głębokiej integracji z wybraną architekturą (kamera, procesor, pamięć), co w przypadku tworzenia platformy rozwojowej może stanowić zbyt duże ograniczenie.

4. Koncepcja implementacji w układzie FPGA

Wykorzystując wiedzę dotyczącą możliwych do zastosowania metod podłączenia kamery do systemu obliczeniowego, postanowiono stworzyć kontroler, będący uniwersalną platformą obliczeń wizyjnych typu *embedded*. Ogólną strukturę systemu przedstawia

rysunek 5. Głównym założeniem koncepcji jest zapewnienie łatwego i uniwersalnego połączenia dwóch kamer CMOS z interfejsem cyfrowym w taki sposób, aby mogły one bezpośrednio transmitować dane obrazowe do jednego (lub dwóch) z ośmiu w pełni niezależnych bloków pamięci.

Przewidziano zastosowanie jednostki obliczeniowej, realizującej wybrane algorytmy przetwarzania obrazów w czasie rzeczywistym, w oparciu o dane pobrane przez kamery. Transfer danych odbywać się może równocześnie z przeprowadzaniem obliczeń na danych zgromadzonych w innych blokach i zapisem wyników cząstkowych i ostatecznych do jeszcze innych bloków.



Rys. 5. System oparty na układzie FPGA
Fig. 5. The proposed embedded system based on FPGA device

Równoległość operacji na tym poziomie jest możliwa do uzyskania w zasadzie jedynie w układzie reprogramowalnym. Część struktury układu stanowi wówczas jednostka DMA, służąca do niezależnego transferu danych na linii kamera-pamięć. Możliwe jest kreowanie struktury kanałów transmisji na dowolnym kierunku. Funkcja ta ma duże znaczenie w przypadku chęci przeprowadzania obliczeń w czasie rzeczywistym, nie tylko o charakterze strumieniowym, ale także tych wymagających znajomości szerokiego kontekstu wybranych fragmentów obrazów. Można zapisywać jeden blok pamięci nowymi danymi w czasie przeprowadzania obliczeń na innym bloku, który zawiera dane poprzednio zapisanej ramki.

Warto zwrócić uwagę na fakt, że blok interfejsu kamery również jest wbudowany w strukturę FPGA, a więc istnieje możliwość zmiany interfejsu na drodze programowej. Dzięki temu można łatwo przystosować system do podłączenia kamer wyposażonych w różne interfejsy. Jednostka przetwarzania obrazów stanowi kolejny blok struktury reprogramowalnej, dzięki czemu może wydajnie realizować szeroką gamę algorytmów przetwarzania obrazów w czasie rzeczywistym, wymagających wykonywania bardzo wielu stosunkowo prostych operacji arytmetycznych i logicznych na dużych blokach danych.

Całość systemu podzielono w fizycznej realizacji na dwie części. Pierwszą stanowi płytka z układem FPGA oraz układami bezpośrednio z nim współpracującymi – jest to płytka 6-warstwowa, o wymiarach 81x97 mm. Znajduje się na niej układ reprogramowalny Xilinx Spartan 6, osiem pamięci statycznych RAM, ośmiobitowych, o pojemności 2 MB każda, a także trzy generatory zegarowe i pamięci konfiguracyjne FPGA (zawierające *bitstream* wygenerowany dzięki opisowi w języku VHDL). Zastosowanie pamięci statycznych zamiast dynamicznych może znacząco przyspieszyć realizację wielu algorytmów, wymagających częstego dostępu do wielu często zmieniających

lokacji adresowych. Nie są wówczas konieczne zaawansowane mechanizmy optymalizacyjne, w tym tworzenie dodatkowych buforów typu *cache*.

Druga z kolei płytka to płytka bazowa, o wymiarach 106x97mm. Wymiary są tak dobrane, aby możliwe było zabudowanie jej w obudowie stanowiącej bazę dla kamer, złącz i elementów wykonawczych (silniki i przekładnie). Na płycie umieszczono, oprócz złącza do podłączenia pierwszej płytki z FPGA, układy zasilające, a także peryferyjne, stanowiące fizyczny sprzęg z kamerami (bufory zmieniające poziomy napięcie i odpowiednie złącza), wyświetlaczem, modułami komunikacyjnymi (dla standardów USB, RS485 oraz Ethernet) i wykonawczymi (złącze dla dedykowanego kontrolera silników krokowych). Przewidziano również złącze dla modułu mikroprocesorowego, dla realizacji zadań bezpośrednio niezwiązanych z przetwarzaniem obrazów.

Rozdzielenie strefy obliczeniowej od interfejsowej jest wygodne, gdyż warstwy logiczne wbudowane w układ FPGA są z definicji rekonfigurowalne, a więc wbudowanie warstw fizycznych na tej samej płycie do układu FPGA stanowiłoby dużą przeszkodę w ewolucji projektu.

5. Podsumowanie

Zastosowanie któregośkolwiek z omówionych rozwiązań zależy ściśle od konkretnych wymagań stawianych przed budowanym systemem. W większości zastosowań celowe jest wykorzystanie standardowych urządzeń pracujących w powszechnie przyjętych standardach, zarówno kamery, jak i platformy obliczeniowej. Gdy jednak konieczna jest budowa własnego urządzenia typu *embedded*, wykorzystującego w swojej konstrukcji kamerę, korzystnie jest zbudować własny interfejs kamerasystem, dopasowany do przyjętej architektury i maksymalnie wykorzystujący możliwości sprzętowe detektora obrazu. Zaproponowany typ oparty o układ FPGA doskonale spełnia to założenie, będąc rozwiązaniem bardzo uniwersalnym, umożliwiającym zastosowanie wielu typów kamer. Ponadto rozbudowany system przesyłu danych między kamerami a pamięciami stanowić może wydajną platformę mogącą służyć do realizacji nawet złożonych algorytmów przetwarzania obrazów.

Projekt finansowany z Ministerstwa Nauki i Szkolnictwa Wyższego nr 0128/R/100/2010/12.

6. Literatura

- [1] Web-i Smart Inspection Camera, Webview product specification, <http://www.webinspection.com/Web-i%20Datasheet.pdf>
- [2] Gorgoń M.: Przetwarzanie obrazów wysokiej rozdzielczości w układach FPGA, PAK (7), pp.752-754, 2010.
- [3] Smart camera manufacturers links <http://www.smartcamera.it/links.htm>
- [4] http://www.dalsa.com/corp/markets/ccd_vs_cmos.aspx
- [5] Omnivision OV14810 image sensor, <http://www.ovt.com/products/category.php?id=14>
- [6] ST Microelectronics VS6524 image sensor, <http://www.st.com/stonline/products/families/imaging/camera/vs6524.htm>
- [7] AT91SAM9XE512 ARM9 microcontroller, http://www.atmel.com/dyn/products/product_card_v2.asp?part_id=4263
- [8] Herrington D.: Easy Image Processing: Camera Interfacing for Robotics, http://www.atmel.com/dyn/resources/prod_documents/issue4_pg39_43_Robotics.pdf
- [9] Zawadzki A.: Lighting Fitting Controller Using Image Processing System, Preprints of IFAC Workshop on Programmable Devices and Embedded Systems, PDES 2009, ISSN 1474-6670.