



# PORTFOLIO

(na wniosek grodzkiej spółdzielni mieszkaniowej w Nowym Sączu)

## Opracowanie koncepcji identyfikacji zdarzeń na podstawie danych z systemu monitoringu

*Autor: dr Dariusz Jamróż*

Centrum Inteligentnych Systemów Informatycznych Akademia Górniczo-Hutnicza im. Stanisława Staszica al. Mickiewicza 30, 30-059 Kraków  
budynek C-2 pokój 426 tel.: 12 617 44 53 www.isi.agh.edu.pl [isi@agh.edu.pl](mailto:isi@agh.edu.pl)



## 1. MOTYWACJA

Opracowanie powstało w wyniku zgłoszenia zapotrzebowania na system rozpoznawania zdarzeń, zgłoszone przez Grodzką Spółdzielnię Mieszkaniową w Nowym Sączu. Potrzeba powstania takiego systemu wynika z faktu, iż Zarząd Spółdzielni jako jeden z priorytetów uznaje zapewnienie ochrony mienia oraz zapewnienie bezpieczeństwa osób przebywających na terenach zarządzanych przez Spółdzielnię. W tym celu postanowiono wykorzystać nowoczesne środki techniczne do automatycznej identyfikacji niebezpiecznych sytuacji oraz poinformowania odpowiednich służb mogących we właściwy sposób, z jak najmniejszym opóźnieniem zareagować na powstałe zdarzenia.

## 2. ZAŁOŻENIA WSTĘPNE

Ponieważ tereny obsługiwane przez Spółdzielnię stanowią stosunkowo duży obszar, Zarząd Spółdzielni określił, że docelowa liczba użytych kamer koniecznych do monitorowania chronionego obszaru to 280 kamer, każda o rozdzielczości 5 megapikseli, przesyłająca 5 klatek na sekundę.

## 3. RODZAJE ROZPOZNAWANYCH ZDARZEŃ

Identyfikowane przez system powinny być wszelkie zdarzenia, stanowiące zagrożenie bezpieczeństwa ludzi oraz zagrożenie ochrony mienia. Poniżej przedstawiono przykłady zdarzeń, które system powinien zidentyfikować. Rysunki 1 oraz 2 stanowią dwie klatki filmu otrzymanego z jednej z kamer znajdujących się na osiedlu zarządzanym przez Grodzką Spółdzielnię Mieszkaniową. Przedstawiają zdarzenie, w którym młody człowiek ucieka przed właścicielem jednego z lokali.



Rysunek 1. Klatka filmu uzyskanego z kamery monitoringu, przedstawiająca zdarzenie, w którym młody człowiek ucieka przed właścicielem jednego z lokali. Kadr pochodzi z filmu udostępnionego przez Grodzką Spółdzielnię Mieszkaniową w Nowym Sączu.



Rysunek 2. Sytuacja z rysunku 1 po upływie 1/4 sekundy.

Rysunki 3, 4 stanowią dwie klatki filmu przedstawiającego zdarzenie, w którym samochód wjeżdża tyłem z otwartymi drzwiami w obszar, w którym nie powinien się znaleźć (tereny zielone).



Rysunek 3. Samochód wjeżdża tyłem z otwartymi drzwiami w obszar zielony. Kadr pochodzi z filmu udostępnionego przez Grodzką Spółdzielnię Mieszkaniową w Nowym Sączu.



Rysunek 4. Dalszy przebieg sytuacji z rysunku 3.

## 4. ALGORYTMY

W praktyce do rozpoznawania zdarzeń na podstawie danych uzyskanych z kamery stosuje się następujące metody:

- proste algorytmy rozpoznające zmiany tła (zmiany wydzielonych fragmentów tła, uznanych za obszary krytyczne)
- rozpoznające konkretne kształty (np. walizki, pakunki, ludzi, samochody, zwierzęta)
- rozpoznające wektory przemieszczenia (kształtów lub zmian tła)
- rozpoznające ogólne zdarzenia na podstawie poprzednich algorytmów

### 4.1 Algorytmy rozpoznające zmiany tła

Jest to klasa najprostszych w implementacji a jednocześnie w określonych sytuacjach skutecznych metod. Zasada działania polega na wyznaczeniu obszarów kadru (obrazu przekazywanego z kamery), które w normalnej sytuacji nie powinny się zmieniać. Zmiana wyglądu tych obszarów świadczy o naruszeniu obszaru chronionego i jest identyfikowane jako zdarzenie powodujące zadziałanie alarmu. Ze względu na swoją prostotę, często stosowana jest w prostych systemach alarmowych.

### 4.2 Algorytmy rozpoznające konkretne kształty

Rozpoznawanie kształtów jest znacznie bardziej skomplikowane niż rozpoznawanie zmian tła. Służą one np. do rozpoznawania sylwetek ludzi, zwierząt, pojazdów, walizek, plecaków. Jest wiele metod służących takiemu rozpoznawaniu, zaczynając od prostych algorytmów sprawdzających wielkość i proporcje po zastosowanie np. 96 wejściowych sieci neuronowych, analizujących 96 analizowanych cech.

### 4.3 Algorytmy rozpoznające wektory przemieszczenia

Badany jest tutaj dodatkowo sposób, w jaki zmienia się obraz w czasie. Rozpoznając np. sylwetki ludzi, możemy stwierdzić analizując kolejne klatki filmu przesłanego przez kamerę, prędkość i kierunek, w którym te sylwetki się przemieszczają.

### 4.4 Rozpoznawanie zdarzeń

Stosując wyżej wymienione algorytmy, można identyfikować bardziej złożone zdarzenia.

Przykłady zdarzeń rozpoznawanych na podstawie wektorów ruchu postaci:

- szybkie przemieszczanie się w określonym kierunku kilku postaci ludzkich (ucieczka)
- pojawienie się wielu postaci ludzkich (zgromadzenie)
- szybki, zmienny ruch kilku postaci ludzkich (bójka)

Przykłady zdarzeń rozpoznawanych na podstawie wektorów ruchu pojazdu:

- przemieszczanie się w niewłaściwym kierunku (jazda „pod prąd”)

- zatrzymanie samochodu w miejscu chronionym przed zatrzymaniem (parkowanie w miejscu niedozwolonym, blokowanie przejazdu)



Rysunek 5. Nałożone na siebie rysunki 1 oraz 2 po rozpoznaniu sylwetek ludzi.



Rysunek 6. Nałożone na siebie rysunki 1 oraz 2 po wyznaczeniu wektorów ruchu, uzyskanych na podstawie rozpoznania sylwetek ludzi na obu rysunkach.

Na podstawie analizy wektorów ruchów możliwe jest również:

- śledzenie obiektów,
- liczenie obiektów,
- detekcja pozostawionych obiektów,

Dla przykładu, na rysunku 5 przedstawiono nałożone na siebie rysunki 1 oraz 2 po rozpoznaniu sylwetek ludzi. Na podstawie rozpoznanych sylwetek oraz zmianie ich położenia, uzyskano wektory ruchu rozpoznanych sylwetek (rysunek 6). Wektory takie przy uwzględnieniu różnicy czasu oraz perspektywy pozwalają uzyskać wektory prędkości przemieszczenia. Na podstawie takich wektorów możemy stwierdzić sposób zachowania się ludzi, np. to że człowiek lub grupa ludzi biegnie w określonym kierunku.

## 5. SPRZĘT

Rozpoznawanie obrazów jako dziedzina informatyki od początku swojego istnienia zgłaszało ogromne zapotrzebowanie na moc obliczeniową, konieczną do przeprowadzenia prawidłowej identyfikacji. Związane jest to zarówno z algorytmami o dużej złożoności obliczeniowej jak i z ogromną ilością przetwarzanej informacji, koniecznej do przeprowadzenia procesu rozpoznawania. Tym bardziej, przy rozpoznawaniu zdarzeń na podstawie monitoringu uzyskanego z 280 kamer poważną przeszkodą staje się moc obliczeniowa konieczna do przetwarzania tak dużej ilości danych. Należy zwrócić uwagę na fakt, że kamera działająca z rozdzielczością 5 megapikseli, przy 24 bitowym kolorze, przesyłająca 5 klatek na sekundę generuje strumień 75 megabajtów na sekundę. Z 280 kamer powstaje więc łączny strumień danych o wielkości 21 gigabajtów danych w ciągu jednej sekundy, na podstawie których należy rozpoznać wystąpienie zdefiniowanych zdarzeń. W praktyce, przy mniejszych rozdzielczościach, przy pomocy procesora znajdującego się w jednym komputerze klasy PC jesteśmy w stanie analizować dane z kilku, maksymalnie kilkunastu kamer. W związku z powyższym, w analizowanej sytuacji możliwe są cztery podstawowe podejścia:

- użycie 28 komputerów PC, każdy obsługuje 10 kamer. Wyniki przekazywane do komputera, w którym zbierane są wszystkie informacje o zagrożeniach,
- wykorzystanie mocy obliczeniowej akceleratorów graficznych (obliczenia scentralizowane),
- wykorzystanie układów FPGA do przeprowadzenia obliczeń (obliczenia rozproszone),
- skorzystanie z gotowych rozwiązań sprzętowo programowych.

**Centrum Inteligentnych Systemów Informatycznych** Akademia Górniczo-Hutnicza im. Stanisława Staszica al. Mickiewicza 30, 30-059 Kraków  
budynek C-2 pokój 426 tel.: 12 617 44 53 www.isi.agh.edu.pl [isi@agh.edu.pl](mailto:isi@agh.edu.pl)

### 5.1. Użycie 28 komputerów PC, każdy wykorzystuje 4 rdzeniowy procesor główny

Każdy z komputerów przeprowadza obliczenia związane z identyfikacją niebezpiecznych sytuacji dla 10 kamer. Następnie wyniki poprzez sieć zbierane są w komputerze znajdującym się w centrali. W przypadku rozpoznania sytuacji niebezpiecznej, komputer ten zgłasza operatorowi miejsce i rodzaj zagrożenia oraz przekazuje widok z odpowiedniej kamery.

Przy takiej konstrukcji należy liczyć się ze znacznym poborem mocy. Komputer klasy PC przeprowadzający obliczenia może pobierać 300W każdy. W związku z powyższym cały system złożony z 28 komputerów może pobierać nawet 8.4 kW mocy elektrycznej. W takim przypadku, roczne zużycie energii elektrycznej potrzebnej do ciągłego działania systemu może przekroczyć 73500 kWh.

### 5.2. Wykorzystanie mocy obliczeniowej akceleratorów graficznych

W wyniku ciągłego wzrostu wymagań przez gry komputerowe mocy obliczeniowej kart graficznych, potrzebnej do wyświetlania scen w wysokiej rozdzielczości z uwzględnieniem wielu szczegółów modelowanej w grach rzeczywistości, następuje ciągle i znaczny wzrost mocy obliczeniowej produkowanych kart graficznych. Już wiele lat temu moc obliczeniowa kart graficznych znacznie przekroczyła możliwości obliczeniowe procesora głównego komputera klasy PC. Dodatkowo producenci kart graficznych umożliwiają producentom oprogramowania, wykorzystanie mocy obliczeniowych swoich kart do wykonywania dowolnych obliczeń – nawet zupełnie niezwiązanych z grafiką. Przykładem niech będzie biblioteka CUDA, dzięki której możemy pisać programy w języku C, wykonywane na wielu rdzeniach karty graficznej.

Zwiększenie szybkości obliczeń poprzez karty graficzne, możliwe jest jedynie w przypadku wykonywania programów opartych o algorytmy dobrze skalowalne czyli takie, które można skutecznie zrównoleglić. Niestety, nie każdy algorytm pozwala na wykonywanie jednocześnie wielu swoich fragmentów na wielu jednostkach liczących. Na szczęście znaczna większość algorytmów służących do przetwarzania obrazów oraz do rozpoznawania jest dobrze skalowalnych. Z tego powodu zastosowanie akceleratorów graficznych do rozpoznawania obrazów jest jak najbardziej uzasadnione. Tym bardziej, że przyspieszenie działania programów wykonywanych na akceleratorach graficznych w porównaniu z wykonaniem tych samych zadań na procesorze głównym jest znaczne.



W ramach wstępnych badań, została przebadana moc obliczeniowa najnowszej i najszybszej obecnie, popularnej karty graficznej, opartej o procesor graficzny produkowany przez firmę NVidia. Jest to karta graficzna GeForce GTX 780 Ti. Karta ta zawiera 2880 rdzeni służących do prowadzenia obliczeń. Rdzenie liczące w trakcie prowadzonych badań pracowały z częstotliwością 1163 MHz, natomiast pamięć karty GDDR5 o pojemności 3072 MB i szynie 384 bitowej, pracowała z efektywną częstotliwością 7000 MHz. Wykonywano obliczenia służące do przetwarzania wielowymiarowych obiektów. Obliczenia te na jednorrdzeniowym procesorze Intel z zegarem 2.4 MHz wykonywały się przez 80.9 sekundy, na czterordzeniowym procesorze 2.4 MHz te same obliczenia wykonywały się przez 20.5 sekundy oraz na czterordzeniowym procesorze 3.2 GHz wykonywały się w ciągu 15.8 sekundy. Następnie sprawdzono ilość czasu konieczną do wykonania tych samych obliczeń z wykorzystaniem badanego akceleratora graficznego: wykonały się one w czasie 0.26 sekundy. Te same obliczenia wykonały się więc na akceleratorze graficznym ponad 60 razy szybciej niż na czterech rdzeniach procesora 3.2 GHz oraz ponad 311 razy szybciej niż na jednorrdzeniowym procesorze 2.4 GHz. Widać z tego, że przy dobrze skalowanych problemach, wykorzystanie akceleratorów graficznych do przyspieszenia obliczeń jest jak najbardziej uzasadnione. Zestawienie wyników pokazano w tabeli 1.

JEDNOSTKA LICZĄCA	CZAS WYKONANIA	SZYBKOŚĆ WZGLĘDNA
Procesor Intel 2.4GHz	80.9 s	311 razy wolniej niż GTX780Ti
Procesor Intel 2.4GHz, 4 rdzenie	20.5 s	78 razy wolniej GTX780Ti
Procesor Intel 3.2GHz, 4 rdzenie	15.8 s	60 razy wolniej GTX780Ti
GeForce GTX 780 Ti, 2880 rdzeni	0.26 s	1

Tabela 1. Zestawienie szybkości wykonania obliczeń przez procesor główny komputera klasy PC, względem szybkości wykonania tych samych obliczeń przez akcelerator graficzny GeForce GTX 780 Ti (2880 rdzeni liczących, pracujących z zegarem 1163 MHz).

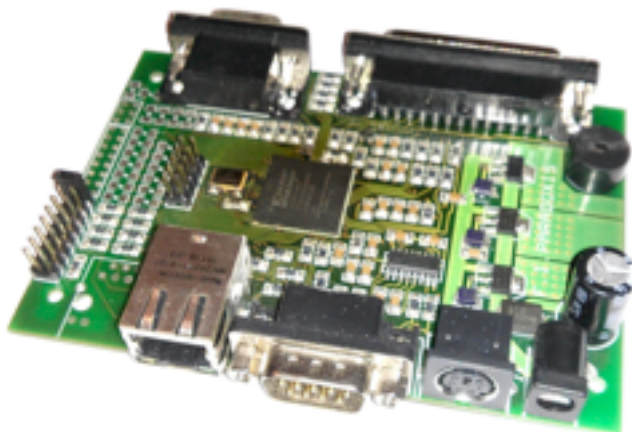
Dzięki zastosowaniu akceleratorów graficznych możliwa jest więc znaczna redukcja liczby komputerów, koniecznych do prowadzenia obliczeń służących do identyfikacji zdarzeń pochodzących z 280 kamer. Dodatkowym problemem jest obsługa odbioru przez system strumienia danych pochodzących z tak wielu kamer. Jeśli przyjęlibyśmy transfer bez kompresji, to kamera działająca z rozdzielczością 5 megapikseli, przy 24 bitowym kolorze, przesyłająca 5 klatek na sekundę generuje strumień 75 megabajtów na sekundę. Z 280 kamer powstanie więc łączny strumień danych o wielkości 21 gigabajtów w przeciągu jednej sekundy. Na szczęście przy zastosowaniu kompresji, można taki łączny strumień znacznie zredukować do 3 gigabitów na sekundę. Dekompresję danych otrzymanych z kamer przed analizą tych danych, również można przeprowadzić przy użyciu akceleratora graficznego.

Należy na końcu wspomnieć, że sam opisany wyżej akcelerator graficzny pobiera 250W a system komputerowy zawierający taki akcelerator nie pobiera więcej niż 600W mocy elektrycznej.

### 5.3. Wykorzystanie układów FPGA do przeprowadzenia obliczeń

Układy FPGA należą do rekonfigurowalnych układów logicznych. Wewnętrzną strukturę logiczną tych układów, można ustalać w sposób dowolny. Dzięki temu można tworzyć specjalizowane procesory, realizujące znacznie szybciej specyficzne działania, potrzebne do realizacji konkretnych algorytmów. Na przykład, jeżeli potrzebujemy analizować cały wiersz obrazu złożony z 1600 punktów, gdzie każdy punkt jest reprezentowany przez 24 bity, to możemy przy pomocy FPGA zrealizować procesor, którego słowo będzie składało się z  $1600 \times 24 = 38400$  bitów. Jedyną przeszkodą może być pojemność struktury FPGA, czyli wartość, która określa liczbę bramek przeliczeniowych z których może się składać tworzony układ. Możemy również definiować bezpośrednio potrzebne, skomplikowane operacje wykonywane na słowach procesora, których realizacja sprzętowa może wykonać się znacznie szybciej niż algorytm korzystający z typowych rozkazów procesora ogólnego przeznaczenia. Do struktury FPGA możemy wpisywać dowolny układ cyfrowy. Możemy konstruować zarówno procesory jak i układy wejścia/wyjścia, układy zarówno kombinacyjne jak i sekwencyjne. Przykład komercyjnego komputera zrealizowanego na układzie FPGA przedstawia Rys.7. Poza układem FPGA zawiera on jedynie stabilizatory napięć oraz układ konwersji napięć.

W strukturze FPGA zawarto w nim zarówno procesor jak i układy wejścia/wyjścia, realizujące sprzętowo protokoły konieczne do transmisji przez sieć UDP/IP, port



równoległy LPT, port szeregowy RS, klawiaturę PS2 oraz kartę graficzną.

Rysunek 7. Przykład komercyjnego komputera zrealizowanego na układzie FPGA (zdjęcie za zgodą producenta: Firmy Informatycznej Paradoxis).

Dodatkowym atutem układów FPGA jest fakt, że wewnętrzną strukturę układu można zmieniać w dowolnym momencie, nawet po wlutowaniu do płytki drukowanej. Dzięki temu opłacalna staje się produkcja układów specjalizowanych nawet dla małych serii produkcyjnych. Dlatego można wykorzystać te układy do konstrukcji specjalizowanych urządzeń, potrafiących identyfikować zdefiniowane zdarzenia na podstawie obrazu z kamery. Każdej z 280 kamer analizowanego systemu monitoringu można przyporządkować jeden taki specjalizowany układ. Po rozpoznaniu określonego zdarzenia będzie on wysyłał informację do komputera głównego o zaistnieniu takiej sytuacji. W przypadku, gdy dane otrzymane z kamery poddane zostały kompresji, również dekompresję danych można przeprowadzić przy użyciu FPGA.

Jednym z głównych producentów układów FPGA jest firma Xilinx. Ostatnia, opracowana rodzina układów to Virtex-7, Kintex-7 i Artix-7. Dodatkowo dwa pierwsze występują również w wersji Ultra Scale. Różnią się one wielkością zasobów udostępnianych przez te układy. Możliwe jest również zastosowanie układów starszych generacji. Wybór konkretnego układu zależy od zapotrzebowania projektowanego urządzenia na zasoby zawarte w takich układach. Należy go dokonać dopiero po opracowaniu i wyborze algorytmów oraz zamianie ich na układ elektroniczny możliwy do wpisania do FPGA. Wtedy dopiero można określić, na jaką strukturę jest rzeczywiste zapotrzebowanie. Moc elektryczna zużywana przez takie układy przeważnie nie przekracza kilku watów. Przy dużym obciążeniu większych struktur działających z dużymi częstotliwościami, moc ta nie przekracza dwudziestu watów.

#### 5.4. Gotowe rozwiązania

Obecnie wiele ośrodków badawczych prowadzi badania związane z automatycznym rozpoznawaniem zagrożeń, przy użyciu systemów monitoringu. Przykładem znanej na całym świecie instytucji, mogącej poszczycić się ogromną liczbą wynalazków i patentów, prowadzącej szeroko zakrojone badania w zakresie automatycznego rozpoznawania na podstawie monitoringu może być amerykańska HRL LABORATORIES, LLC ([www.hrl.com](http://www.hrl.com)). Przykładem firmy proponującej gotowe, komercyjne rozwiązania w tym zakresie, może być angielska firma Miragex Ltd ([miragex.co.uk](http://miragex.co.uk)).

## 6. WNIOSKI

Realizacja analizowanego systemu identyfikacji zdarzeń jest problemem złożonym, w skład którego wchodzi trzy podstawowe zagadnienia:

- 1) Precyzyjne zdefiniowanie identyfikowanych zdarzeń.
- 2) Opracowanie, wybór i implementacja algorytmów służących rozpoznawaniu zdarzeń.
- 3) Realizacja odpowiedniego modelu sprzętowego, pozwalającego sprawnie działać systemowi identyfikacji na podstawie monitoringu złożonego z 280 kamer.