

Przemysław MAZUREK<sup>1</sup>  
Krzysztof OKARMA<sup>1</sup>

## **ALGORYTMY SUPERROZDZIELCZOŚCI W INTELIGENTNYCH SYSTEMACH TRANSPORTOWYCH DO ŚLEDZENIA ODLEGŁYCH POJAZDÓW**

*W artykule przedstawiono możliwości zastosowania algorytmów superrozdzielczości do poprawy rozdzielczości obrazu pojazdów w ruchu znajdujących się w znacznej odległości od kamery, a zarazem przy kamerze umiejscowionej na niewielkiej wysokości nad drogą. Zaproponowane rozwiązanie bazuje na estymacji tła i dodatkowym wykorzystaniu operatora morfologicznego dylatacji w celu uzyskania lepszej jakości wynikowego obrazu zrekonstruowanego na podstawie sekwencji kolejnych klatek wideo.*

## **SUPERRESOLUTION ALGORITHMS FOR DISTANCE VEHICLES TRACKING IN INTELLIGENT TRANSPORT SYSTEMS**

*The article presents the usage possibilities for the superresolution algorithms for the resolution increase of the image with vehicles in motion located in a considerable distance from the camera while the camera is situated low over the road. The proposed solution is based on the background estimation and the additional use of the morphological dilation operator to obtain a better quality of the resulting image reconstructed from a sequence of video frames.*

### **1. WSTĘP**

Inteligentne Systemy Transportowe (ITS – ang. *Intelligent Transport Systems*) są jedną z wiodących metod poprawy jakości transportu, głównie samochodowego. Zapewniają one automatyzację procesów sterowania ruchem i jego monitoring w czasie rzeczywistym, co daje możliwość adaptacyjnego pośredniego sterowania pojazdami. Systemy takie składają się z części wykonawczej (podsystem sterujący, w skład którego wchodzi m.in. sygnalizacja świetlna, znaki o zmiennej treści itp.), części pomiarowej (w jej skład wchodzi różnego typu czujniki drogowe) oraz algorytmów sterowania wraz ze sprzętową infrastrukturą teleinformatyczną.

---

<sup>1</sup> Wyższa Szkoła Techniczno-Ekonomiczna w Szczecinie, Wydział Transportu Samochodowego; 71-244 Szczecin; ul. Klonowica 14. Tel: +48 91 424-08-75, Fax: +48 91 424-08-76 E-mail: mazurek @wste.szczecin.pl, okarma @wste.szczecin.pl

Główne źródło kosztów w ITS stanowi system pomiarowy, który pozwala na pomiar różnych wielkości związanych z pojazdami w ruchu. Istnieje bardzo duża liczba różnorodnych czujników pomiarowych, jednak nie ma idealnego czujnika mierzącego wszystkie wartości, co oznacza konieczność zastosowania czujników różnego rodzaju dostosowanych do oczekiwań operatora ITS, jak również metod fuzji danych.

W przypadku zaawansowanych ITS wbudowane czujniki drogowe zlokalizowane są np. co 100 metrów na każdym z pasów ruchu. Alternatywnym rozwiązaniem jest wykorzystanie bardziej zaawansowanych algorytmów estymacji większej liczby parametrów. Rozwiązanie to jest możliwe dla wielu rodzajów czujników np. światłowodowe czujniki nacisku mogą nie tylko zliczać ilość i prędkość pojazdów na pasie ruchu, ale również umożliwiać estymację ciężaru pojazdów [3].

Podobna sytuacja ma miejsce w systemach wizyjnych, w których zliczanie pojazdów i wyznaczanie prędkości są jedynie podstawowymi zadaniami, a poprzez wykorzystanie lepszych algorytmów i dzięki zmniejszeniu się cen jednostek przetwarzania o dużej wydajności obliczeniowej możliwe jest dalsze rozszerzenie możliwości pomiarowych. W konsekwencji możliwe jest usprawnienie istniejących systemów bez inwestycji w rozbudowę systemów pomiarowych.

Wadą systemów wizyjnych jest ich rozdzielczość wynikająca z geometrii sceny. Jedynie w najprostszych systemach dedykowanych wyłącznie do celów rozpoznawania tablic rejestracyjnych kamera jest pochylona pod znacznym kątem z ograniczonym zakresem widoczności drogi.

W wypadku śledzenia pojazdów [1] kamery montować można albo na bardzo wysokich słupach lub budynkach, co jest praktycznie ograniczone do centrów niektórych większych miast. Bardziej typowym rozwiązaniem jest instalowanie kamer na wysokości kilku metrów, co niestety ogranicza możliwości pomiarowe przy zastosowaniu rozwiązań konwencjonalnych. Z tego też powodu zasięg pracy kamer jest stosunkowo niewielki i ogranicza się do kilkudziesięciu metrów. Za pomocą bardziej zaawansowanych metod algorytmicznych można natomiast estymować parametry pojazdów z niewielkiej wysokości kamery.

Niezależnie od wysokości istnieje kolejny, równie ważny, parametr ograniczający widoczność kamery, którym jest rozdzielczość zarówno matrycy jak i układu optycznego. Innym istotnym czynnikiem ograniczającym jest widoczność, zależna od stanu atmosfery, temperatury, zapylenia, opadów atmosferycznych, pory dnia i pory roku.

Czynniki wpływające na ograniczenie rozdzielczości dla większej odległości od kamery są istotnym utrudnieniem funkcjonowania systemów wizyjnych. Ważnymi zaletami systemów wideo są bezkontaktowy pomiar, brak konieczności ingerencji w nawierzchnie oraz relatywnie atrakcyjna cena, ponieważ jedna kamera może zazwyczaj zastąpić kilkanaście czy nawet kilkadziesiąt czujników montowanych w drodze.

## **2. SUPERROZDZIELCZOŚĆ W SYSTEMACH WIDEO**

Zwiększenie rzeczywistej rozdzielczości obrazu możliwe jest na kilka sposobów. Pierwszy z nich polega na zwiększeniu liczby kamer (tego samego typu). Poprzez wykorzystanie algorytmów superrozdzielczości [2] można uzyskać wówczas większą rzeczywistą rozdzielczość obrazu. Rozwiązanie to nie jest najlepsze z punktu widzenia

kosztów systemu, ponieważ wymaga większej liczby kamer, a końcowy rezultat i tak jest zależny od wielu czynników takich jak np. położenie wzajemne kamer.

Drugim rozwiązaniem jest zainstalowanie pojedynczej kamery w taki sposób, aby nie była ona sztywno zamocowana, co pozwalałoby na jej nieznaczne poruszanie się. Jest to jednakże rozwiązanie dość kłopotliwe w praktycznej realizacji, choć pozwala ono na wykorzystanie kilku klatek obrazu do uzyskania jednej o wyższej rozdzielczości również za pomocą algorytmów superrozdzielczości.

Trzecim rozwiązaniem jest wykorzystanie kamery statycznej do rejestracji wielu klatek obrazu. Możliwe jest w takim wypadku uzyskanie obrazu o większej rozdzielczości za pomocą algorytmów superrozdzielczości dla sceny, która przesuwa się względem kamery. Rozwiązanie to jest najprostsze i nie wymaga rozbudowy czy modyfikacji istniejącej infrastruktury akwizycji obrazu i możliwe jest do realizacji na drodze wyłącznie programowej.

Algorytmy superrozdzielczości pozwalają na zwiększenie rozdzielczości wyłącznie w wypadku, gdy występuje większa ilość klatek obrazu zarejestrowanych dla tej samej sceny, jednak nie mogą być one uzyskane z tego samego położenia kamery. Nawet małe przesunięcie (nie tylko o pojedyncze piksele, ale także o ich ułamkowe części - przesunięcia subpikselowe) pozwala zgromadzić więcej informacji z obrazu, które mogą być wykorzystane do syntezy obrazu lepszej jakości o większej rozdzielczości. Nie jest to jednak klasyczna technika interpolacji obrazu z wykorzystaniem jednej klatki obrazu, dla której „sztucznie” wyliczane są wartości brakujących pikseli – interpolacja taka nie zwiększa ilości informacji w obrazie, w przeciwieństwie do superrozdzielczości.

Konieczne przesunięcie pomiędzy sceną a kamerą uzyskać można dla pojazdów znajdujących się w ruchu. Ponieważ tło jest statyczne (nie zmienia swojego położenia względem kamery) w celu prawidłowego funkcjonowania technik superrozdzielczości konieczna jest jego eliminacja. Nie jest to operacja wpływająca negatywnie na ilość użytecznej informacji w obrazie, a wręcz przeciwnie. W procesie estymacji parametrów, np. zliczania pojazdów, wyznaczania prędkości, itp., określenie, które obszary obrazu odpowiadają za tło (np. powierzchnię drogi, drzewa, budynki) jest zadaniem prawie zawsze wykonywanym w celu redukcji ilości obliczeń i poprawy skuteczności algorytmów detekcji.

W celu wyznaczenia obszarów, w których znajdują się pojazdy w ruchu oraz tło dla każdej klatki obrazu można wykorzystać jeden z wielu algorytmów estymacji tła. Najprostsze z nich bazują na uśrednianiu obrazów wielu klatek zgodnie z zależnością:

$$\hat{E} = \frac{1}{N} \sum_{i=1}^N E_i \quad (1)$$

gdzie:

$\hat{E}$  - estymowane tło,

$E_i$  - obraz zarejestrowany o numerze  $i$  ( $i$ -ta klatka sekwencji wideo).

Przykładowy rezultat uśredniania ilustruje rysunek 1.



Rys. 1 Przykładowe klatki obrazu (górne trzy rzędy) i uśredniony obraz (estymowane tło) z 10000 klatek.

Możliwe jest także zastosowanie rozwiązania rekurencyjnego bazującego na wygładzającym filtrze wykładniczym (ang. *exponential smoothing filter*) zgodnie ze wzorem:

$$\hat{E}_n = \alpha \hat{E}_{n-1} + (1 - \alpha) E_n \quad (2)$$

gdzie:

$\hat{E}_n$  - estymowane tło,

$\hat{E}_{n-1}$  - tło estymowane w poprzednim kroku,

$E_n$  - nowy zarejestrowany obraz ,

$\alpha$  - współczynnik wygładzania z zakresu (0-1).

Dla efektywnego wykorzystania wygładzania wykładniczego współczynnik  $\alpha$  powinien być bliski jedności, ale nieco mniejszy od niego, np. 0.99.

Innym algorytmem estymacji jest algorytm porównujący dwie klatki ze sobą – miejsca, w których są takie same wartości pikseli przez określony dłuższy odcinek czasu klasyfikowane są jako tło, natomiast zmienne fragmenty obrazu uznawane są za reprezentacje obiektów ruchomych (w naszym przypadku pojazdów). Nowe piksele, które zostały sklasyfikowane jako tło, są uśredniane wraz z poprzednimi wartościami tła w sposób standardowy (1) lub wykładniczy (2) w celu aktualizacji tła.



Rys. 2 Przykład selekcji obrazu pojazdu (górny wiersz) zwykłą maską (środkowy wiersz z lewej) oraz maską z dylatacją (środkowy wiersz z prawej).

Ekstrakcja obrazu pojazdów z aktualnie zarejestrowanego obrazu możliwa jest, gdy dysponujemy obrazem tła. Dla każdego piksela porównywane są wszystkie wartości koloru między obrazem tła i aktualnie zarejestrowanym. Jeśli różnica między nimi jest mniejsza dla wszystkich kolorów od założonej tolerancji uznać można, iż piksel reprezentuje tło, w przeciwnym przypadku jest to piksel obiektu ruchomego, co można wyrazić zależnością:

$$M_i = \begin{cases} 0 & : \left( |R_i - \hat{R}_i| < \varepsilon \right) \cap \left( |G_i - \hat{G}_i| < \varepsilon \right) \cap \left( |B_i - \hat{B}_i| < \varepsilon \right) \\ 1 & : \text{w przeciwnym przypadku} \end{cases} \quad (3)$$

gdzie:

- $M$  - maska,
- $i$  - numer piksela,
- $\varepsilon$  - tolerancja,
- $R_i, G_i, B_i$  - wartości kanałów RGB koloru zarejestrowanego obrazu,
- $\hat{R}_i, \hat{G}_i, \hat{B}_i$  - wartości kanałów RGB koloru estymowanego uprzednio tła.

Algorytm ten pozwala na wyznaczenie maski binarnej  $M$ . Maską tą w zależności od wartości tolerancji jest także zaszumiona. Ze względu na fakt, iż każda kamera posiada zawsze szumy zależne od temperatury i warunków oświetlenia (czas ekspozycji), przyjęta wartość tolerancji  $\varepsilon$  nie może być równa zero, a jedynie bliska zera. Powoduje to sytuację taką, iż część pikseli reprezentujących pojazd (zwłaszcza na granicy obszaru tło-pojazd) zostaje przydzielona do tła, co oznacza, że maska jest zwykle mniejsza niż obraz pojazdu.

W przypadku małych rozmiarów śledzonych obiektów jest to nieakceptowalne, ponieważ w efekcie względny rozmiar obiektów byłby istotnie zmniejszony.

W celu redukcji tego efektu można wykorzystać operację morfologiczną dylatacji (ang. *dilation*) polegającą na powiększeniu zadanego obszaru obrazu binarnego (maski) o wszystkie piksele bezpośrednio z nim sąsiadujące [6]. Ponieważ dylatacja powoduje rozszerzenie obszaru o wartościach 1 na sąsiednie piksele, wpływ tolerancji i szumów jest zredukowany. Ponadto rozszerzenie analizowanego obszaru poza obrys pojazdu na fragment drogi nie wpływa ujemnie na wynik pracy algorytmu superrozdzielczości, jedynie zwiększa nieznacznie jego złożoność obliczeniową (większa liczba analizowanych pikseli).

Ponieważ operacja pasowania w algorytmie superrozdzielczości jest iteracyjna, ważne jest zachowanie odpowiedniej kolejności obrazów w celu uzyskania najlepszego wyniku. Wynika to z faktu, iż aktualny rezultat pasowania jest łączony z kolejnym nowym obrazem. W związku z tym jeśli w tracie zarejestrowanej sekwencji obrazów występuje zmiana skali obiektów (np. zbliżający się pojazd), nie należy rozpoczynać pasowania od klatki zawierającej najmniejszy obiekt, ale od tej, gdzie jest on możliwie największy.

Dla przypadku oddalania się pojazdu od kamery to pierwszy obraz jest „najbliższy”, jednakże może być on niekompletny (np. może zawierać fragment widocznego pojazdu). W celu zwiększenia uniwersalności proponowanego algorytmu zdecydowano się na rozwiązanie bazujące na kolejności naprzemienną poczynając od obrazu środkowego, a następnie dodawaniu kolejnych klatek poprzednich i następnych naprzemiennie zgodnie z regułą:

$$v = m + (-1)^{m+n} \left\lfloor \frac{n+1}{2} \right\rfloor \quad (5)$$

gdzie:

$n = 1, 2, 3, \dots, N$

$m$  – numer obrazu początkowego, od którego rozpoczyna się proces pasowania w algorytmie superrozdzielczości,

$v$  – numer kolejnego obrazu dostarczanego algorytmowi superrozdzielczości,

$N$  – liczba obrazów dodawanych do początkowego,

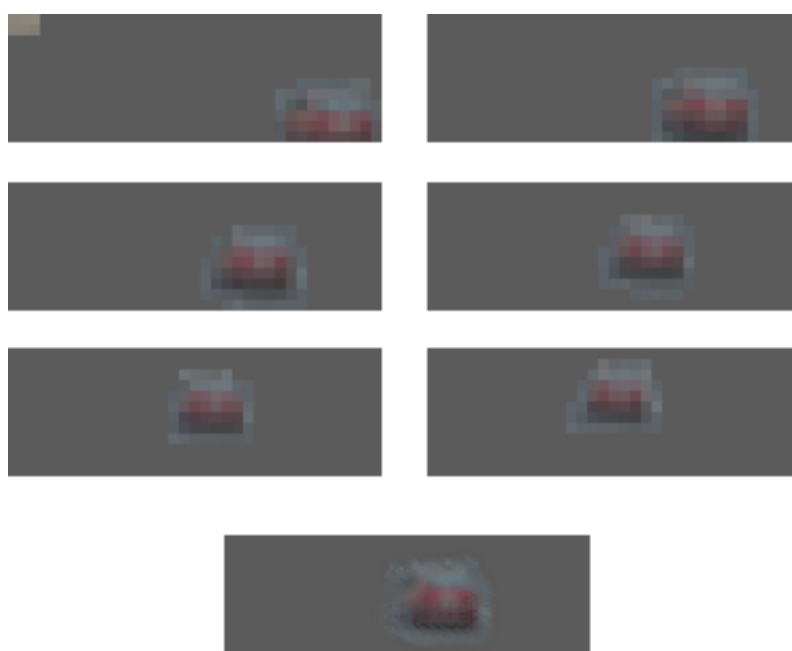
$\lfloor \cdot \rfloor$  – operator *floor* (zaokrąglenia w dół do najbliższej liczby całkowitej)

Wprowadzone obrazy pojazdów mają oczywiście różne rozmiary z uwagi na zmianę odległości od kamery. Proces pasowania bazuje w takim wypadku albo na kosztownym obliczeniowo badaniu wszystkich kombinacji pasowania lub na stochastycznym pasowaniu - w tym wypadku wykorzystano algorytm pasowania losowego (metoda Monte Carlo). Jeśli stopień dopasowania nowego obrazu do już połączonych jest zbyt mały, wtedy obraz ten jest odrzucany jako potencjalnie pogarszający jakość wynikowego obrazu.

W wypadku pasowania obrazu konieczne jest maksymalne ograniczenie stopni swobody algorytmu w celu poprawy jakości i zarazem skrócenia czasu obliczeń. Z tego też powodu dopuszczalne są operacje translacji (przemieszczenia) oraz skalowania obrazu w określonych granicach, jednak niedopuszczalna jest rotacja obrazu. W wypadku rejestracji serii obrazów w ciągu kilku sekund nie występują zwykle zmiany oświetlenia,

można więc założyć stałe warunki ekspozycji, co znacznie upraszcza proces pasowania, ponieważ zmiany jasności obrazów nie muszą być analizowane przez algorytm superrozdzielczości.

Bardzo ważnym elementem procedury jest zawężenie obszaru poszukiwań parametrów dla obrazów, tak aby wyniki były wiarygodne i niezależne od zawartości sceny. Ich wybór jest zadaniem jednoznacznym dla pojazdów znajdujących się w ruchu w przybliżeniu prostoliniowym np. jadącym na jednym pasie ruchu. W przypadku skrętu pojazdu (np. na skrzyżowaniu) zadanie to jest bardzo złożone.



Rys. 3 Przykładowe klatki obrazu (wybrane z 60 użytych przez algorytm superrozdzielczości - górne wiersze) oraz wynik pracy algorytmu (dolny wiersz)



Rys. 4 Oryginalny obraz o wysokiej rozdzielczości

Dla większej liczby pojazdów w ruchu w krótkich odcinkach czasu położenie pojazdów jest relatywnie stałe, jedynie dla obszarów bezpośrednio przed i za skrzyżowaniami występują nieliniowości znacznie zakłócające pracę algorytmów. Z tego też powodu zaproponowane podejście może być stosowane przede wszystkim na dłuższych odcinkach dróg bez skrzyżowań. Wykorzystanie kamer i algorytmów superrozdzielczości jest w tym wypadku właściwe i celowe, ponieważ na takich odcinkach instalacja złożonej i drogiej infrastruktury innego typu nie jest ekonomicznie uzasadniona.

W celu weryfikacji pracy algorytmu superrozdzielczości i algorytmów wspomagających wykorzystano obrazy o znacznie większej rozdzielczości i zmniejszono ich rozdzielczość na potrzeby testów, co pozwala na porównanie wyników w sposób wiarygodny.

Warto zauważyć iż algorytm superrozdzielczości oparty na metodzie Monte Carlo ma zmienny czas działania, co jest niewątpliwie jego wadą. Dla obrazów, w których nie wykorzystano dylatacji proces obliczenia obrazu wynikowego trwał często powyżej kilku godzin z uwagi na problem ze znalezieniem właściwego dopasowania. Powiększenie analizowanego obszaru za pomocą dylatacji pomimo zwiększonej złożoności pojedynczej iteracji, pozwoliło na zredukowanie całkowitego czasu wyznaczenia obrazu docelowego do ułamka sekundy.

### 3. WNIOSKI

Zaprezentowana metoda poprawy jakości obrazów oparta na technikach superrozdzielczości może być z powodzeniem wykorzystana do śledzenia odległych obiektów, co ma duże znaczenie w Inteligentnych Systemach Transportowych. Uzyskane wyniki pozwalają na zwiększenie skuteczności rozpoznawania i klasyfikacji pojazdów przy wykorzystaniu informacji dostępnych z pojedynczej kamery zlokalizowanej nawet w znacznej odległości od drogi.

### 4. BIBLIOGRAFIA

- [1] Blackman S., Popoli R.: *Design and Analysis of Modern Tracking Systems*. Artech House Radar Library, August 1999.
- [2] Irani M., Peleg S.: *Super Resolution from Image Sequences*. Proceedings of the 10th International Conference on Pattern Recognition, pp. 115-120, June 1990.
- [3] Klein L.A.: *Sensor Technologies and Data Requirements for ITS*. Norwood, Massachusetts, Artech House ITS library 2001.
- [4] Okarma K. Mazurek P.: *Techniki super rozdzielczości w systemach identyfikacji tablic rejestracyjnych pojazdów*. Drogi – lądowe-powietrzne-wodne, no. 4/2009 pp. 62-73
- [5] Okarma K. Mazurek P.: *The Analysis of the Super-Resolution Techniques for the Automatic Vertical Road Signs Recognition Systems*. Materiały XIII Międzynarodowej Konferencji „Komputerowe Systemy Wspomagania Nauki, Przemysłu i Transportu – TransComp 2009”
- [6] Tadeusiewicz R., Korohoda P.: *Komputerowa analiza i przetwarzanie obrazów*. Wydawnictwo Fundacji Postępu Telekomunikacji, Kraków 1997.

### 5. PODZIĘKOWANIA

Artykuł powstał dzięki wsparciu w ramach grantu Ministerstwa Nauki i Szkolnictwa Wyższego nr N509 399136 „Estymacja trajektorii ruchu pojazdów z wykorzystaniem analizy bayesowskiej oraz algorytmów cyfrowego przetwarzania obrazów“.