

Krzysztof OKARMA¹
Przemysław MAZUREK¹

AUTOMATYCZNA OCENA JAKOŚCI OBRAZÓW UZYSKIWANYCH TECHNIKAMI SUPERROZDZIELCZOŚCI DO CELÓW ROZPOZNAWANIA TABLIC REJESTRACYJNYCH

W artykule przeanalizowano możliwość wykorzystania nowoczesnych porównawczych miar jakości obrazów cyfrowych do automatycznej oceny efektów działania algorytmów superrozdzielczości w aspekcie rozpoznawania tablic rejestracyjnych pojazdów. Pomimo konieczności znajomości oryginalnego obrazu o wysokiej rozdzielczości, uzyskane wyniki pozwalają poczynić założenia dotyczące wyboru możliwie najefektywniejszej techniki superrozdzielczości do celów późniejszego rozpoznawania numerów rejestracyjnych.

AUTOMATIC IMAGE QUALITY ASSESSMENT OF THE SUPER-RESOLUTION IMAGES FOR THE REGISTER PLATE NUMBER RECOGNITION

The article examines the possibility of using modern full-reference digital image quality assessment methods to evaluate automatically the effects of super-resolution algorithms in the aspect of vehicle's register plate recognition. Despite having knowledge of the original high-resolution image, the results obtained allow to make some assumptions about the possible selection of the most effective super-resolution techniques for further recognition of registration numbers.

1. WSTĘP

Techniki superrozdzielczości [4] stanowią jedną z najbardziej przydatnych grup algorytmów cyfrowego przetwarzania obrazów z punktu widzenia zastosowań w Inteligentnych Systemach Transportowych. Mają one szczególne znaczenie w wypadku obserwacji pojazdów ze znacznej odległości, co ogranicza w istotny sposób rozdzielczość użytecznego fragmentu obrazu, często do zaledwie kilkunastu lub kilkadziesiąt pikseli. W takiej sytuacji dzięki zastosowaniu technik superrozdzielczości możliwe jest śledzenie toru ruchu pojazdów nawet reprezentowanych przez kilka pikseli. Algorytmy superrozdzielczości znajdują również zastosowanie przy rozpoznawaniu znaków

¹ 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: okarma@wste.szczecin.pl, mazurek@wste.szczecin.pl

drogowych w systemach wspomagających kierowców [10], jak również przy rozpoznawaniu numerów tablic rejestracyjnych [9]. Możliwe jest także ich wykorzystanie jako technik wspomagających np. do klasyfikacji marek pojazdów na podstawie analizy obrazu przodu lub tyłu pojazdu, co może stanowić informację uzupełniającą w wypadku np. częściowego przesłonięcia tablicy rejestracyjnej (np. przez śnieg). W takiej sytuacji pożądane jest wykorzystanie bazy danych pojazdów, dzięki której możliwe byłoby szybkie wyszukanie pojazdów spełniających zadane kryteria np. marka (określona technikami wizyjnymi na podstawie „znaczką” na masce), kolor (łatwy do wyznaczenia ze zdjęcia) i np. początek numeru rejestracyjnego. Ze względu na fakt, iż istnieje wiele odmian technik super-rozdzielczości, które nie zawsze spełniają oczekiwania w konkretnych zastosowaniach, konieczny jest wybór zarówno konkretnego algorytmu dedykowanego do określonego problemu, jak też jego parametrów. Pomocną techniką może w tym wypadku okazać się, dynamicznie rozwijająca się w ostatnich latach, automatyczna ocena jakości obrazów cyfrowych.

2. METODY OCENY JAKOŚCI OBRAZÓW

Zagadnienie oceny jakości obrazów cyfrowych przez wiele lat sprowadzało się do wykorzystania błędu średniokwadratowego (ang. *Mean Squared Error* – MSE), zazwyczaj obliczanego dla pary obrazów w skali szarości, z których jeden stanowił obraz referencyjny o „doskonałej” jakości, nie zawierający żadnych zakłóceń ani zniekształceń [1,2]. Na bazie MSE powstało wiele miar pochodnych, takich jak np. szczytowy stosunek sygnału do szumu (ang. *Peak Signal to Noise Ratio* – PSNR), jednak wszystkie te miary charakteryzują się podobnymi wadami dotyczącymi m.in. czułości na wzajemne przesunięcie obrazów lub ich obrót. W takiej sytuacji nawet obraz o doskonałej jakości, lecz przesunięty zaledwie o jeden lub kilka pikseli względem oryginału byłby oceniony jako zakłócony.

Innym podejściem do zagadnienia oceny jakości obrazów, znacznie bardziej pożądanym w wielu szeroko pojętych zastosowaniach przemysłowych, w tym Inteligentnych Systemach Transportowych, jest tzw. „ślepa” ocena jakości obrazów [6], która nie wymaga znajomości obrazu oryginalnego. Takie metody oceny jakości wydają się szczególnie atrakcyjne w wielu zastosowaniach, gdzie nie ma dostępu do obrazu referencyjnego. Niestety metody „ślepej” oceny jakości charakteryzują się małą uniwersalnością i są zazwyczaj czułe tylko na jeden lub dwa wybrane rodzaje zniekształceń. Pewnym kompromisem pomiędzy podejściem „full-reference” a „no-reference” jest grupa metod korzystających jedynie z częściowej informacji o obrazie oryginalnym (tzw. „reduced-reference”), jednak mają one zasadniczo te same wady, co metody „ślepe”.

Począwszy od roku 2002, który uznać można za przełomowy dzięki propozycji tzw. uniwersalnego wskaźnika jakości obrazu (*Universal Image Quality Index*) [13], należącego do grupy metod porównawczych, nastąpił znaczny postęp w dziedzinie oceny jakości obrazów zgodnej z ludzką percepcją, a co z tym się pośrednio wiąże, również w dużej mierze skorelowanej ze skutecznością automatycznego rozpoznawania obrazów zawierających zakłócenia. Spośród kilkunastu nowych miar jakości, zazwyczaj weryfikowanych poprzez badanie korelacji z ocenami subiektywnymi dla grupy obrazów testowych, największą popularność uzyskała propozycja tzw. podobieństwa strukturalnego (ang. *Structural Similarity* – SSIM) [14] stanowiącego rozwinięcie uniwersalnego wskaźnika jakości obrazu. W metodzie tej, czulej na trzy najbardziej typowe rodzaje

zniekształceń tj. zmianę jasności, zmianę kontrastu oraz zaburzenie struktury obrazu, następuje skanowanie obrazu referencyjnego i ocenianego na zasadzie przesuwnego okna (zazwyczaj jest to okno Gaussa o rozmiarze 11×11 pikseli), dla którego wyliczany jest lokalny wskaźnik jakości fragmentu obrazu zgodnie z zależnością (po uproszczeniu):

$$SSIM = \frac{(2 \cdot \overline{xy} + C_1) \cdot (2 \cdot \sigma_{xy} + C_2)}{(\overline{x}^2 + \overline{y}^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (1)$$

gdzie: $\overline{x}$ – wartość średnia obrazu referencyjnego wewnątrz okna

$\overline{y}$ – wartość średnia obrazu ocenianego wewnątrz okna

σ_x^2 – wariancja obrazu referencyjnego wewnątrz okna

σ_y^2 – wariancja obrazu ocenianego wewnątrz okna

σ_{xy} – kowariancja

C_1, C_2 – stałe zabezpieczające przed niestabilnością

Wzór (1) w postaci rozwiniętej stanowi iloczyn trzech czynników odpowiadającym trzem wymienionym rodzajom zniekształceń. Rolą stałych zabezpieczających przed niestabilnością jest zapobieganie dzieleniu przez wartości zerowe, co mogłoby nastąpić dla bardzo ciemnych oraz „płaskich” (o stałym kolorze) fragmentów obrazu. Zalecane wartości są równe $C_1 = 0.01 \times L$ oraz $C_2 = 0.03 \times L$, gdzie L jest maksymalnym dopuszczalnym poziomem jasności w obrazie (zwykle $L = 255$).

Przesuwając okno co jeden piksel uzyskiwana jest mapa jakości obrazu, która po uśrednieniu pozwala uzyskać skalarny wskaźnik jakości z zakresu $<-1;1>$.

3. TECHNIKI SUPERROZDZIELCZOŚCI UŻYTE W EKSPERYMENTACH

W celu weryfikacji efektów uzyskiwanych za pomocą technik superrozdzielczości wykorzystane zostały dwa wskaźniki jakości: klasyczny błąd średniokwadratowy (MSE) obliczany dla obrazów znormalizowanych (o jasności z zakresu $<0;1>$) oraz podobieństwo strukturalne (SSIM). Eksperymenty przeprowadzono na szeregu obrazów, z których dwa uznane za reprezentacyjne zaprezentowano w pracy. Dla każdego z obrazów wygenerowano po 10 obrazów o 4-krotnie zmniejszonej rozdzielczości z losowo wybranymi przesunięciami i obrotami, które zostały wykorzystane jako dane wejściowe do poszczególnych algorytmów superrozdzielczości.

Biorąc pod uwagę fakt, iż wynik działania algorytmu superrozdzielczości uzależniony jest zarówno od użytej metody estymacji wektorów ruchu dla poszczególnych obrazów, jak również wykorzystanego algorytmu rekonstrukcji, przeanalizowano kombinacje czterech metod estymacji oraz kilku technik rekonstrukcji.

Każdy algorytm superrozdzielczości wymaga obrazów wejściowych zarejestrowanych z nieznacznymi różniącymi się od siebie położeniami kamery. Przesunięcia pomiędzy pikselami mogą być nawet ułamkowe, co pozytywnie wpływa na jakość uzyskanego obrazu końcowego. Pierwszy etap obliczeń stanowi dopasowywanie obrazów (estymacja wektorów ruchu). Jedną z metod estymacji, zaproponowaną przez Vandewalle’a [12], wykorzystuje fakt, iż przesunięcie obrazu jest również widoczne w częstotliwościowej

charakterystyce fazowej uzyskiwanej w wyniku Dyskretnej Transformacji Fouriera (DFT). Z kolei obrót obrazu jest widoczny w częstotliwościowej charakterystyce amplitudowej, co pozwala na estymację obu tych parametrów przy dodatkowym odfiltrowaniu wysokich częstotliwości, w których może zaistnieć zjawisko aliasingu. Inna metoda oparta na DFT została zaproponowana przez Marcela [8]. Z kolei Lucchese i Cortelazzo [7] zaproponowali wykorzystanie kąta nachylenia linii przecinającej zero widocznej w widmie obrazu do celu estymacji obrotu oraz korelacji fazowej do wyznaczenia przesunięć. Innym podejściem [5] jest wykorzystanie rozwinięcia w szereg Taylora dla obrazów o czterokrotnie i dwukrotnie obniżonej rozdzielczości oraz dla obrazu oryginalnego.

Rekonstrukcja obrazu również może przebiegać na kilka sposobów, najbardziej typowym z nich jest interpolacja bikubiczna, często wykorzystywana przy powiększaniu obrazów cyfrowych. Inne podejście bazuje na rzutowaniu zbiorów wypukłych (ang. *Projection onto Convex Sets* - POCS), w którym następuje iteracyjne rzutowanie pikseli na siatkę o podwyższonej rozdzielczości i odcięcie wysokich częstotliwości. Proces ten jest powtarzany aż do osiągnięcia zbieżności (metoda Papoulisa-Gerchberga). W standardowej metodzie POCS zamiast odcięcia wysokich częstotliwości następuje filtracja dolnoprzepustowa za pomocą filtru aproksymującego funkcję PSF (ang. *Point Spread Function*) będącą odpowiedzią impulsową zogniskowanej kamery.

Iteracyjny algorytm wstecznej propagacji [3] bazuje z kolei na zgrubnej estymacji obrazu o wysokiej rozdzielczości, po czym następuje iteracyjne dodanie obrazu „gradientowego” stanowiącego sumę błędów pomiędzy wszystkimi obrazami wejściowymi a obrazem wynikowym. Wyliczenie gradientu jako mediany wszystkich błędów zamiast ich sumy prowadzi natomiast to tzw. algorytmu *Robust Super Resolution* [15].

Ostatnią metodą użytą w eksperymentach, jest tzw. *Structure-Adaptive Normalized Convolution* (SANC) [11] oparta na znormalizowanym splocie z opcjonalnym dodatkowym przebiegiem mającym na celu wyostrenie obrazu za pomocą filtrów Gaussa o adaptacyjnie zmienianym rozmiarze i orientacji.

Uzyskane w efekcie obrazy oraz wartości wskaźnika SSIM i znormalizowanego błędu średniokwadratowego pokazano w tabelach 1 i 2, natomiast rysunki 1 i 2 przedstawiają oryginalne obrazy oraz obrazy wejściowe o zmniejszonej rozdzielczości.

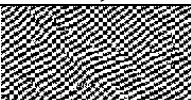
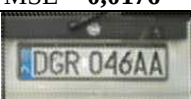
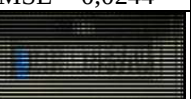
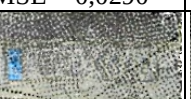
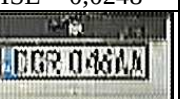
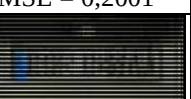
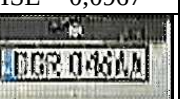


Rys.1. Oryginalny obraz testowy 1 (150x90 pikseli) wraz z 10 obrazami o 4-krotnie zmniejszonej rozdzielczości użytymi do rekonstrukcji



Rys.2. Oryginalny obraz testowy 2 (150x90 pikseli) wraz z 10 obrazami o 4-krotnie zmniejszonej rozdzielczości użytymi do rekonstrukcji

Tab. 1. Wyniki uzyskane dla obrazu testowego 1

Metoda estymacji ruchu Algorytm rekonstrukcji	Vandewalle / Süsstrunk / Vetterli [12]	Marcel / Briot / Murrieta [8]	Lucchese / Cortelazzo [7]	Keren / Peleg / Brada [5]
Interpolacja bikubiczna	 SSIM = 0,6261 MSE = 0,0228	 SSIM = 0,4399 MSE = 0,0336	 SSIM = 0,1153 MSE = 0,0475	 SSIM = 0,3729 MSE = 0,0370
Papoulis- Gerchberg	 SSIM = 0,4542 MSE = 0,0195	 SSIM = 0,0556 MSE = 0,2059	 SSIM = 0,1519 MSE = 0,0687	 SSIM = 0,0259 MSE = 0,2830
Propagacja wsteczna [3]	 SSIM = 0,0046 MSE = 0,3039	 SSIM = 0,2999 MSE = 0,0372	 SSIM = 0,0423 MSE = 0,1258	 SSIM = 0,2320 MSE = 0,0384
Robust SR [15]	 SSIM = 0,4641 MSE = 0,0279	 SSIM = 0,3053 MSE = 0,0400	 SSIM = 0,2466 MSE = 0,0387	 SSIM = 0,2195 MSE = 0,0488
POCS	 SSIM = 0,6138 MSE = 0,0176	 SSIM = 0,4783 MSE = 0,0244	 SSIM = 0,2140 MSE = 0,0290	 SSIM = 0,4620 MSE = 0,0248
Structure- Adaptive Normalized Convolution 1-pass [11]	 SSIM = 0,6394 MSE = 0,0185	 SSIM = 0,0526 MSE = 0,2001	 SSIM = 0,1161 MSE = 0,0454	 SSIM = 0,0819 MSE = 0,0967
Structure- Adaptive Normalized Convolution 2-pass [11]	 SSIM = 0,6595 MSE = 0,0179	 SSIM = 0,0506 MSE = 0,2029	 SSIM = 0,1180 MSE = 0,0449	 SSIM = 0,0699 MSE = 0,1002

Tab. 2. Wyniki uzyskane dla obrazu testowego 2

Metoda estymacji ruchu Algorytm rekonstrukcji	Vandewalle / Süsstrunk / Vetterli [12]	Marcel / Briot / Murrieta [8]	Lucchese / Cortelazzo [7]	Keren / Peleg / Brada [5]
Interpolacja	 SSIM = 0,7006 MSE = 0,0119	 SSIM = 0,3589 MSE = 0,0252	 SSIM = 0,1728 MSE = 0,0359	 SSIM = 0,7538 MSE = 0,0087
Papoulis- Gerchberg	 SSIM = 0,5542 MSE = 0,0149	 SSIM = 0,0700 MSE = 0,1003	 SSIM = 0,0610 MSE = 0,0854	 SSIM = 0,1743 MSE = 0,0543
Propagacja wsteczna [3]	 SSIM = 0,0048 MSE = 0,3139	 SSIM = 0,1992 MSE = 0,0403	 SSIM = 0,0053 MSE = 0,3127	 SSIM = 0,0056 MSE = 0,3150
Robust SR [15]	 SSIM = 0,0252 MSE = 0,2146	 SSIM = 0,1777 MSE = 0,0504	 SSIM = 0,0257 MSE = 0,1562	 SSIM = 0,0516 MSE = 0,1471
POCS	 SSIM = 0,6707 MSE = 0,0119	 SSIM = 0,4757 MSE = 0,0187	 SSIM = 0,2856 MSE = 0,0308	 SSIM = 0,6717 MSE = 0,0118
Structure- Adaptive Normalized Convolution 1-pass [11]	 SSIM = 0,6992 MSE = 0,0119	 SSIM = 0,0640 MSE = 0,0991	 SSIM = 0,1844 MSE = 0,0326	 SSIM = 0,7522 MSE = 0,0091
Structure- Adaptive Normalized Convolution 2-pass [11]	 SSIM = 0,7055 MSE = 0,0124	 SSIM = 0,0607 MSE = 0,1013	 SSIM = 0,1835 MSE = 0,0328	 SSIM = 0,7421 MSE = 0,0108

4. ANALIZA WYNIKÓW I WNIOSKI

Analizując uzyskane wyniki można stwierdzić większą przydatność wskaźnika podobieństwa strukturalnego (SSIM) do oceny jakości obrazów uzyskiwanych za pomocą różnych algorytmów superrozdzielczości w porównaniu z klasycznym błędem średniokwadratowym. W tabelach 1 oraz 2 zaznaczono pogrubieniem najwyższe wartości tego wskaźnika, podobnie jak najmniejsze wartości błędu średniokwadratowego. Jak można zauważyć, nie wszystkie metody estymacji ruchu i rekonstrukcji prowadzą do akceptowalnych wyników, a szczególnie niedoskonała okazała się metoda wstecznej propagacji, a także jej modyfikacja - *Robust SR*.

Można zauważyć, iż w kilku wypadkach wysokie wartości SSIM, a szczególnie mały błąd MSE, są spowodowane głównie przez zgodność tła, a nie czytelność numeru rejestracyjnego. Dobrze widoczne jest to np. w wypadku obrazu nr 2 dla metody POCS z estymacją ruchu metodą Marcela. Interesujące jest również porównanie wyników rekonstrukcji metodami Papoulisa-Gerchberga i POCS dla obrazu nr 1 z estymacją ruchu metodą Vandewalle'a, dla których mały błąd średniokwadratowy lepiej świadczy o dobrej czytelności numeru rejestracyjnego aniżeli wartość SSIM. Z kolei przy interpolacji bikubicznej dla tego samego wypadku przy stosunkowo dużej wartości błędu MSE uzyskana jest wysoka wartość podobieństwa strukturalnego i dobra czytelność tablic.

Dokładność pasowania obrazów uzależniona jest również od tła tablicy rejestracyjnej, o czym można się przekonać analizując wyniki uzyskane dla obu obrazów przy estymacji ruchu metodą Kerena, Pelega i Bragi. Interesującym kierunkiem dalszych badań wydaje się zatem określenie warunków, w których w celu uzyskania wyższej skuteczności działania algorytmów superrozdzielczości należy dokonać ekstrakcji obrazu samej tablicy rejestracyjnej, tak aby pasowanie następowało wyłącznie na podstawie kształtu znaków bez wpływu pozostałych części pojazdu widocznych na obrazie.

5. BIBLIOGRAFIA

- [1] Eskicioglu A., Fisher P., Chen S.: *Image Quality Measures and Their Performance*. IEEE Transactions on Communications, vol. 43 no. 12, pp. 2959-2965, 1995
- [2] Eskicioglu A.: *Quality Measurement for Monochrome Compressed Images in the Past 25 Years*. Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing, Istanbul, Turkey 2000, pp. 1907-1910.
- [3] Irani M., Peleg S.: *Improving Resolution by Image Registration*. Graphical Models and Image Processing, vol. 53, pp. 231-239, 1991.
- [4] Irani M., Peleg S.: *Super Resolution from Image Sequences*. Proceedings of the 10th International Conference on Pattern Recognition, pp. 115-120, June 1990.
- [5] Keren D., Peleg S., Brada R.: *Image Sequence Enhancement Using Sub-Pixel Displacement*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 742-746, June 1988
- [6] Li X.: *Blind Image Quality Assessment*. Proceedings of the IEEE International Conference on Image Processing, Rochester, USA, pp. 449-452, 2002
- [7] Lucchese L., Cortelazzo G. M.: *A Noise-Robust Frequency Domain Technique for Estimating Planar Roto-Translations*. IEEE Transactions on Signal Processing, vol. 48, no. 6, pp. 1769-1786, June 2000

- [8] Marcel B., Briot M., Murrieta R.: *Calcul de Translation et Rotation par la Transformation de Fourier.*, Traitement du Signal, vol. 14, no. 2, pp. 135-149, 1997
- [9] Okarma K. Mazurek P.: *Techniki super rozdzielczości w systemach identyfikacji tablic rejestracyjnych pojazdów.* Drogi – lądowe-powietrzne-wodne, no. 4/2009 pp. 62-73
- [10] 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”
- [11] Pham T.Q., van Vliet L.J., Schutte K.: *Robust Fusion of Irregularly Sampled Data Using Adaptive Normalized Convolution*, EURASIP Journal on Applied Signal Processing, vol. 2006, Article ID 83268, 12 pages, 2006
- [12] Vandewalle P., Süsstrunk S., Vetterli M., *A Frequency Domain Approach to Registration of Aliased Images with Application to Super-Resolution*, EURASIP Journal on Applied Signal Processing (special issue on Super-resolution), vol. 2006, Article ID 71459, 14 pages, 2006
- [13] Wang, Z., Bovik, A.: *A Universal Image Quality Index*. IEEE Signal Processing Letters, vol. 9, no. 3, pp. 81–84, 2002
- [14] Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: *Image Quality Assessment: From Error Measurement to Structural Similarity*. IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600–612, 2004
- [15] Zomet A., Rav-Acha A., Peleg S.: *Robust Super-Resolution*. Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. 645-650, 2001.

6. 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”.