

# Szybki algorytm dopasowania obrazów dla potrzeb fuzji w czasie rzeczywistym

Marcin Kondej, Barbara Putz, Michał Bartyś

Instytut Automatyki i Robotyki Politechniki Warszawskiej

W artykule zaproponowano prosty, szybki i odporny algorytm dopasowania obrazów, który został zaprojektowany do fuzji obrazów realizowanej w czasie rzeczywistym. Algorytm zaprezentowano na tle znanych rozwiązań, dla których może stanowić interesującą alternatywę. Przyjęto, że dopasowanie dotyczyć będzie zwłaszcza obrazów pozyskiwanych synchronicznie przez kamery TV i IR. Omówiono wyniki testowania kilku wariantów prezentowanego algorytmu i przykłady zastosowań, w tym także w odniesieniu do robotyki.

**Słowa kluczowe:** dopasowanie obrazów, fuzja obrazów, algorytm krawędziowy

Fuzja obrazów jest techniką, która pozwala nie tylko na łączenie obrazów, ale przede wszystkim na wydobywanie ich cech szczególnych oraz na pozyskiwanie jakościowo nowych informacji. Dla zilustrowania zalet tej techniki posłużono się przykładem, który dotyczy ważnego aspektu bezpieczeństwa ruchu lotniczego.

Na rys. 1 przedstawiono obraz samolotu pasażerskiego lecącego nisko nad linią horyzontu. Obraz ten jest pozyskany z ka-



Rys. 1. Obraz samolotu w kamerze światła widzialnego

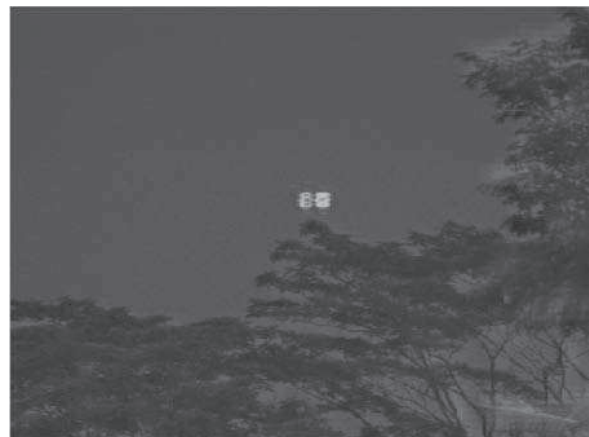
Fig. 1. Image of a plane captured by TV camera



Rys. 2. Obraz samolotu w kamerze termowizyjnej

Fig. 2. Image of a plane captured by IR camera

mery światła widzialnego (TV). Ze względu na panujące warunki atmosferyczne (zamglenie) samolot na tym obrazie jest słabo widoczny dla kontrolera lotów. Na rys. 2 przedstawiono obraz tego samego samolotu pozyskany z kamery termowizyjnej (IR). Ze względu na występujący kontrast termiczny, na obrazie dominującym obiektem są rozgrzane silniki i strugi spalin samolotu. Na obrazie tym nie jest w ogóle widoczny zarys samolotu, a plan pierwszy jest wyraźnie rozmyty. Na rys. 3 pokazano wynik fuzji obu obrazów. Jak łatwo zauważyć, samolot (a ściślej jego silniki) jest bardzo dobrze rozróżnialny na tle dobrze widocznego pierwszego planu. Z punktu widzenia kontrolera lotów obraz ten niesie więcej istotnych dla bezpieczeństwa lotu informacji niż każdy obraz z osobna z przedstawionych na rys. 1 i 2.



Rys. 3. Obraz samolotu po dokonaniu fuzji obrazów z rys. 1 i 2

Fig. 3. Image of a plane as a result of fusion of images from Fig. 1 and Fig. 2

## Założenia

Fuzję obrazów można w uproszczeniu przedstawić jako dwuetapowy proces złożony z fazy dopasowania obrazów i z fazy wykonania fuzji właściwej. Jeśli zostanie zaniedbany efekt zniekształceń obrazów pochodzący z dystorsji układów optycznych kamer, to w fazie dopasowania konieczne jest określenie odpowiednich współczynników definiujących wzajemne relacje geometryczne obu obrazów (a więc współczynników ska-

li, translacji i rotacji). Dopiero po określeniu tych współczynników możliwa jest realizacja fuzji właściwej, czyli ekstrakcja cech i wzajemne nałożenie obrazów.

Realizacja zadania dopasowania nie jest ani trywialna, ani prosta obliczeniowo. Złożoność obliczeniowa tego procesu była i jest w dalszym ciągu główną przeszkodą w implementacji algorytmów przeznaczonych do realizacji dopasowania i fuzji obrazów w czasie rzeczywistym. W artykule zaproponowano prosty, szybki i odporny algorytm dopasowania obrazów, który może stanowić interesującą alternatywę w stosunku do znanych metod. W konstrukcji prezentowanego algorytmu przyjęto następujące założenia ogólne:

- dopasowanie dotyczy dwóch obrazów przedstawiających tę samą scenę
- osie optyczne kamer, z których pochodzą obrazy, są w przybliżeniu równoległe
- rozdzielczości obrazowe obu kamer mogą być różne
- obrazy mogą różnić się przesunięciem i skalą
- efekt dystorsji i rotacji obrazów jest pomijalny
- złożoność obliczeniowa algorytmu powinna być na tyle mała, aby mógł być implementowalny w czasie rzeczywistym przy wykorzystaniu środków współczesnej techniki obliczeniowej.

## Przegląd rozwiązań

Zagadnienie dopasowania obrazów, przy spełnieniu wcześniej sformułowanych założeń, rozwiązywane jest metodami, które można podzielić na trzy grupy.

1. **Klasyczne metody obszarowe**, zwane też metodami dopasowania wzorców, działające w przestrzeni obrazu i wymagające wyznaczenia maksimum miary dopasowania, co zwykle wykonuje się metodami optymalizacji, np. wielowymiarową metodą Powella z użyciem jednowymiarowej metody Brenta [5], metodą Gausa-Newtona, Levenberga-Marquardta [7] lub innymi. Można tu wyróżnić dwa typowe rodzaje metod:

- bazujące na wyznaczeniu korelacji wzajemnej między obrazami [2, 7]. Są to metody o wysokiej złożoności obliczeniowej. Polegają na poszukiwaniu takiej pary okien wyznaczonych na każdym z obrazów, dla której występuje maksymalna miara podobieństwa, liczona najczęściej jako znormalizowana korelacja wzajemna. Metody tego typu są obciążone różnymi wadami, mogą doprowadzić do wyszukania maksimum lokalnego zamiast globalnego, na przykład do dopasowania do siebie obrazów niezawierających żadnej informacji. W najbardziej podstawowej wersji nie sprawdzają się w przypadku dopasowywania obrazów multimodalnych (pochodzących z różnych sensorów)
- maksymalizacji informacji wzajemnej, bazujące na wyznaczeniu entropii, ukierunkowane na dopasowanie obrazów multimodalnych (na przykład medycznych) [5, 7]. Informacja wzajemna jest miarą statystycznej zależności między dwoma zbiorami danych określonymi przez zmienne losowe o znanych rozkładach prawdopodobieństwa. W obliczeniach entropii określającej niepewność jednej zmiennej przy założeniu wiedzy o drugiej zmiennej wykorzystuje się znormalizowane histogramy nakładających się na siebie fragmentów obu obrazów.

2. **Metody bazujące na przekształceniu Fouriera**, działające w dziedzinie częstotliwości obrazu [7], znacznie szybsze niż metody klasyczne. Najbardziej typowa metoda, znana jako metoda Fouriera-Mellina [1, 7], wykorzystuje:

- twierdzenie Fouriera o przesunięciu, co pozwala wyznaczyć przesunięcie dopasowanych obrazów metodą korelacji fazowej, przez wyznaczenie piksu funkcji impulsowej będącej transformatą odwrotną Fouriera względem znormalizowanego iloczynu odpowiednich transformat obrazów wejściowych
- właściwość przekształcenia Fouriera w układzie współrzędnych log-polar (logarytm amplitudy, kąt biegunowy fazy widma) [6] polegającą na tym, że rotacje i skalowanie obiektów powodują jedynie translacje widma ich obrazu w układzie log-polar. Wartość kąta obrotu i skali dopasowywanych obrazów wyznacza się analogicznie jak wartość przesunięcia, metodą korelacji fazowej w układzie log-polar.

Metoda jest szybka, uniwersalna, wyznacza parametry dopasowania w stałym czasie, zależnym jedynie od wielkości obrazów. Algorytm wymaga wyznaczenia sześciu transformat FFT, trzech odwrotnych transformat IFFT i kilku innych, mniej złożonych operacji oraz rozwiązania niejednoznaczności, które mogą wystąpić w odniesieniu do kąta obrotu.

3. **Metody najnowsze**, dające szansę na obliczenia w czasie zbliżonym do rzeczywistego. Najbardziej znaczące osiągnięcia odnotowane są w publikacjach grupy Waterfall Solutions [2, 3, 4]. Metoda ostatnio opublikowana [3, 4] polega na szybkim, wstępnym dopasowaniu obu obrazów, a następnie udoskonalaniu dopasowania metodami dokładnymi.

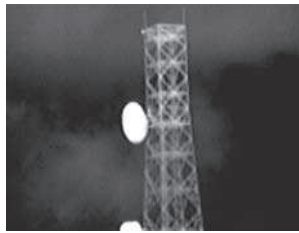
Dopasowanie wstępne ma na celu odnalezienie od razu właściwego maksimum globalnego metodą wyczerpującego poszukiwania i jest najbardziej złożoną częścią obliczeń. Polega ono na podzieleniu obrazu odniesienia na siatkę okien lokalnych i znajdowaniu lokalnego dopasowania do każdego z nich w obrazie badanym z wykorzystaniem miary statystycznej korelacji między obrazami, wprowadzonej w [2]. Otrzymany w ten sposób zbiór dopasowań lokalnych jest następnie organizowany w zbiór spójnych grup dopasowań, dla których znajdowane jest dopasowanie najlepsze; proces ten pozwala zarazem na eliminowanie dopasowań błędnych.

Tak realizowany algorytm jest powtarzany iteracyjnie; każda iteracja bazuje na dopasowaniu uzyskanym w iteracji poprzedniej. Proces wstępnego dopasowania kończy się zwykle po kilku iteracjach, gdy nie można już polepszyć uzyskanego dopasowania. Po nim następuje proces dopasowania dokładnego, z przetwarzaniem bloków obrazu zamiast przesuwanych okien i przy użyciu tej samej, co w procesie wstępnym miary multimodalnego dopasowania. Wykorzystywane jest schemat efektywnej optymalizacji metodą quasi-Newtona, z zastosowaniem, typowej dla wielu metod, piramidy obrazów. Jak podają autorzy, metoda jest bardzo dokładna, odporna na wysoki poziom szumu, obecność uzupełniających cech i występowanie artefaktów. Jest jednocześnie wyjątkowo szybka w porównaniu z metodami klasycznymi, choć na typowych współczesnych komputerach PC nie daje się jeszcze wykonać w czasie rzeczywistym.

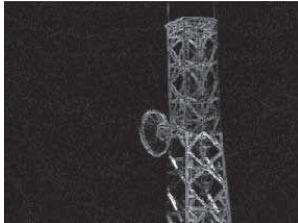
Metoda omówiona w artykule jest wynikiem poszukiwań algorytmu na tyle szybkiego i prostego, by mógł być łatwo zre-



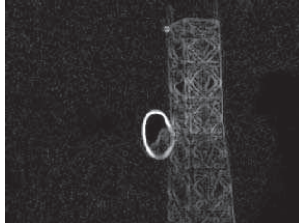
**Rys. 4.** Obraz TV  
**Fig. 4.** TV image



**Rys. 5.** Obraz IR  
**Fig. 5.** IR image



**Rys. 6.** Efekt działania filtra krawędziowego, obraz TV  
**Fig. 6.** Effect of edge filtering, TV image



**Rys. 7.** Efekt działania filtra krawędziowego, obraz IR  
**Fig. 7.** Effect of edge filtering, IR image



**Rys. 8.** Przykład efektu fuzji. Poprawne działanie algorytmu dopasowania  
**Fig. 8.** An example of image fusion. Correct image alignment is visible

alizowany w czasie rzeczywistym (a więc prostszego niż metody z grup 1-3). Pierwsze uzyskane wyniki zdają się potwierdzać realną możliwość osiągnięcia tego celu.

### Szybki algorytm dopasowania obrazów

Szybki algorytm dopasowania obrazów A i B spełniających podane na wstępie założenia polega na wstępnym przeskalowaniu i wyśrodkowaniu obrazu B względem obrazu A – w wyniku tego oba obrazy mają tę samą liczbę wierszy i kolumn – a następnie na iteracyjnej realizacji sekwencji operacji:

1. poddanie obrazu A działaniu filtra krawędziowego
2. poddanie obrazu B działaniu filtra krawędziowego
3. skanowanie odpowiadających sobie kolumn obrazów A i B w celu wyznaczenia takiego wzajemnego ich położenia, dla

którego zachodzi minimum sumy kwadratów różnic ich intensywności (ewentualnie sumy modułów różnic)

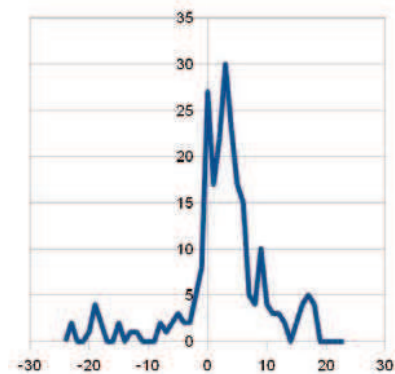
4. wyznaczenie wynikowej translacji  $y_0$  ze zbiorów przesunięć wzajemnych kolumn, z wykorzystaniem metody statystycznej (średnia/mediana/dominanta zbioru)
5. wykonanie translacji obrazu B w kierunku pionowym o wartość  $y_0$
6. skanowanie odpowiadających sobie wierszy obrazów A i B w celu wyznaczenia takiego wzajemnego ich położenia, dla którego zachodzi minimum sumy kwadratów różnic ich intensywności
7. wyznaczenie wynikowej translacji  $x_0$  ze zbiorów przesunięć wzajemnych wierszy
8. wykonanie translacji obrazu B w kierunku poziomym o wartość  $x_0$ .

Iteracja kończy się, gdy kolejny krok nie polepsza już dopasowania (wielkości przesunięć  $x_0$  i  $y_0$  nie ulegają zmianie).

Zastosowane w algorytmie filtry krawędziowe są prostymi filtrami gradientowymi (filtr liczący wartość absolutną różnicy między intensywnością punktu a jego czterema sąsiadami). Dopuszczalne jest jednak użycie filtrów krawędziowych innego typu.

Efekt działania filtrów krawędziowych dla obrazów TV i IR zarejestrowanych synchronicznie (rys. 4 i 5) przedstawiono na rys. 6 i 7.

Jakość uzyskanego dopasowania obrazów była badana przez pomiar miejsca wystąpienia obiektów, ich szerokości i wysokości na obu obrazach. W tym celu wykorzystano ogólnodostępne programy graficzne.



**Rys. 9.** Rozkład częstości występowania translacji poziomych dla obrazów z rys. 4-5 (wartość dominantny = 3, wartość mediany = 4)

**Fig. 9.** Distribution of the horizontal translations for images from Figs. 4-5 (mode value = 3, median value = 4)



Obraz TV



Obraz IR



Fuzja TV + IR

**Rys. 10.** Obrazy testowe z nietypowym źródłem światła (race fosforowe)  
**Fig. 10.** Test images with unconventional light source (phosphorous flare)

Po oszacowaniu odpowiednich współczynników dopasowania sam efekt fuzji, nienależącej do algorytmu dopasowania, nie jest trudny do osiągnięcia (rys. 8). Do celów badawczych użyto mieszania obrazów typu alpha, ze współczynnikiem mieszania równym 50 %. Współczynnik ten może być również obliczany dynamicznie, np. z uwzględnieniem analizy histogramów obu obrazów. Dzięki temu, w warunkach zmiennego oświetlenia lub pogody, możliwe jest przedstawienie fuzji wydobywającej najwyraźniejsze cechy w obserwowanej scenie.

## Dobór parametrów statystycznych

Efektom wyszukiwania najmniejszej sumy bezwzględnych różnic (lub najmniejszej sumy kwadratów różnic) we wszystkich kolumnach i wierszach są zbiory możliwych translacji obrazów w kierunkach pionowym i poziomym (rys. 9). Wartość translacji końcowej w danej iteracji dla pary analizowanych obrazów wyznaczano w wyniku oceny statystycznej populacji uzyskanych translacji. W fazie badań nad algorytmem za wartość tę rozważano przyjęcie:

- średniej arytmetycznej translacji
- mediany zbioru translacji
- dominanty zbioru translacji.

Wszystkie trzy wyżej wymienione parametry statystyczne rozkładu translacji pozwoliły na uzyskanie akceptowalnych wyników. Różnice były jednak widoczne w przypadku szybkości działania algorytmu (liczba wymaganych iteracji) oraz w samej jakości uzyskanego dopasowania obrazów.

Najwięcej czasu na oszacowanie stabilnego (tzn. niezmiennego w kolejnych iteracjach) dopasowania wymagał algorytm oparty na obliczaniu średniej arytmetycznej (średnio pięć iteracji). Zauważono ponadto, że taki algorytm generuje najgorsze dopasowanie, np. w przypadku obserwacji rac fosforowych emitujących znaczne ilości ciepła, widocznych na rys. 10 (błąd rzędu 2 % w skali całego obrazu).

Najlepszą jakość (najmniejszy błąd) dopasowania uzyskano w przypadku zastosowania dominanty. Również szybkość działania była zadowalająca (dwie iteracje). Pośrednie właściwości wykazywał algorytm oparty na medianie. W tym przypadku konieczne były średnio trzy pętle iteracyjne algorytmu. W testach zaprezentowanych w artykule zdecydowano się na wykorzystanie dominanty, również ze względu na jej niewielką złożoność obliczeniową.

## Inne testy algorytmu

Ważną kwestią jest uodpornienie algorytmu na ewentualne błędy. Polega to głównie na eliminowaniu wyników skrajnych ze zbioru otrzymanych wartości translacji. Dodatkowo wydzielane są linie, które nie niosą żadnych informacji. Z analizy wyłączono również obszary położone blisko krawędzi obrazu. Dzięki temu algorytm jest zbieżny i w kilku krokach iteracji poprawnie wyznacza translację, nawet wtedy gdy ilość informacji jest niewielka w obu analizowanych obrazach (rys. 11). Dodatkowo, duże obszary obrazów nieniosące informacji nie wpływają na jakość dopasowania nawet stosunkowo niewielkiego obiektu.



Rys. 11. Efekt fuzji niskoinformacyjnych obrazów testowych

Fig. 11. Result of fusion of „low information” test images

Algorytm jest efektywny także w przypadku występowania obiektów trudno dopasowywalnych w obu obrazach, np. pokazanych na rys. 10. Przeprowadzono również testy na parach obrazów, w których występowały dwa plany wyraźnie różniące się odległością (rys. 12). Na obrazie wynikowym zauważyć można, że pierwszy plan w przeciwieństwie do drugiego nie jest wyraźnie dopasowany. W tym przypadku uwidoczniła się wada algorytmu wynikająca z braku zastosowania w nim złożonej obliczeniowo [3] techniki segmentacji. Translacja jest liczona, lecz dla całego obrazu. W tym przypadku dominujący jest jednak drugi plan obrazu i stąd wynika problem wyznaczenia niedostatecznego dopasowania.

## Podsumowanie

W artykule przedstawiono prosty i szybki w implementacji algorytm dopasowania obrazów o stosunkowo niewielkiej złożoności numerycznej. Algorytm będzie rozwijany i testowany; przewiduje się, że dzięki uzyskanej szybkości będzie się nadał do implementacji w czasie rzeczywistym, w którym dopa-



Obraz TV

Obraz IR

Fuzja TV+IR

Rys. 12. Obrazy testowe z dwoma odległymi planami

Fig. 12. Test images with two distant plans

sowanie obrazów będzie przebiegać współbieżnie z ich pozyskaniem z kamer TV i IR. Zasadniczą wadą algorytmu jest to, że dopasowanie obrazów nie jest doskonałe w pewnych przypadkach szczególnych, gdy w obserwowanej scenie występuje bliski i dalszy plan i to w taki sposób, że powstały efekt stereowizyjny nie jest do pominięcia. Podstawowymi zaletami algorytmu są jego: prostota, szybkość i odporność.

Jako potencjalne obszary zastosowania należy wymienić robotykę oraz sektor wojskowy. Na szczególną uwagę zasługują zastosowania inspekcyjne, gdyż algorytm pozwala na zbudowanie urządzeń służących do kontroli instalacji ciepłowniczych, wymienników ciepła, instalacji elektrycznych czy systemów klimatyzacyjnych.

Algorytm pokazany został w zastosowaniu do fuzji obrazów, która jest typowym, wizualnie przekonującym wykorzystaniem metody dopasowania obrazów. Należy jednak podkreślić uniwersalność algorytmu: może on być stosowany, przy podanych założeniach i ograniczeniach, do wielu innych zagadnień wymagających szybkiego dopasowania obrazów. Przykładem może być nawigacja robota mobilnego w znanym otoczeniu, gdzie trzeba dopasować obraz uzyskany z kamery umieszczonej na robocie z obrazem referencyjnym [1].

## Bibliografia

1. Cassinis, R., Duina, D., Inelli S., Rizzi, A.: *Unsupervised matching of visual landmarks for robotic homing using Fourier-Mellin transform*. Robotics and Autonomous Systems, vol. 40 (2002), 131–138.
2. Heather, J.P., Smith, M.I.: *Multimodal Image Registration with Applications to Image Fusion*. 7<sup>th</sup> International Conference on Information Fusion, vol. 1 (2005).
3. Heather, J.P., Smith, M.I.: *New adaptive algorithms for real-time registration and fusion of multimodal imagery*. Proc. SPIE, vol. 7869 (2010).
4. Heather, J.P., Smith, M.I., Sadler, J., Hickman, D.: *Issues and challenges in the development of a commercialised image fusion system*. Proc. SPIE, vol. 7869 (2010).
5. Maes F., Collignon A., Vandermeulen D., Marchal G., Suetens P.: *Multimodality image registration by maximization of mutual information*, IEEE Transactions on Medical Imaging 16 (1997), 187–198.
6. Reddy B.S., Chatterji B.N.: *An FFT-based technique for translation, rotation and scale-invariant image registration*, IEEE Transactions on Image Processing, vol. 5 (1996), 1266–1271.
7. Zitová B., Flusser J.: *Image registration methods: a survey*. Image and Vision Computing, vol. 21 (2003), 977–1000.

Pracę wykonano częściowo w ramach projektu rozwojowego MNiSW nr O R00 0019 07. ■

## Fast alignment algorithm for real time image fusion

Fast and simple as well as robust image alignment algorithm developed for real time image fusion has been described in this paper. Algorithm has to be considered as an interesting alternative solution compared to the known more complex solutions. It has been assumed, that image alignment will be

suitable particularly for the synchronously acquired images by the TV and IR cameras. The results of the tests of several variants of presented algorithm has been given. Finally, the examples of applications particularly in the area of robotics has been shown.

**Keywords:** image alignment, image fusion, edge algorithm

### inż. Marcin Kondej

Absolwent specjalności Robotyka na Wydziale Mechatroniki Politechniki Warszawskiej. We wrześniu 2010 roku obronił pracę dyplomową inżynierską poświęconą implementacji algorytmu dyskretnej transformacji Fouriera w układzie FPGA. Obecnie na studiach drugiego stopnia tej samej specjalności.



e-mail: [marcin@kondej.eu](mailto:marcin@kondej.eu)

### prof. nzw. dr hab. inż. Barbara Putz

Absolwentka obecnego Wydziału Mechatroniki PW, pracownik Instytutu Automatyki i Robotyki PW. Zajmuje się zagadnieniami zaawansowanego modelowania geometrycznego w zastosowaniach CAD/CAM i grafiki komputerowej, monitorowaniem otoczenia i systemami wizyjnymi w robotyce. Współautorka pierwszego w języku polskim podręcznika z dziedziny modelowania geometrycznego oraz kilku wydań multimedialnych podręczników do nauczania na odległość z dziedziny programowania, algorytmów i struktur danych oraz techniki obrazowej.



e-mail: [b.putz@mchtr.pw.edu.pl](mailto:b.putz@mchtr.pw.edu.pl)

### dr inż. Michał Bartys

Pracownik Instytutu Automatyki i Robotyki Politechniki Warszawskiej. W pracy zawodowej zajmuje się głównie zagadnieniami związanymi z automatyzacją procesów, diagnostyką procesów, systemami tolerowania uszkodzeń, systemami sieciowymi automatyki, inteligentnymi urządzeniami pomiarowymi i wykonawczymi automatyki oraz zastosowaniami logiki rozmytej. Współautor 3 książek i 3 podręczników, autor 101 publikacji w czasopiśmie naukowych i konferencjach, autor i współautor 4 patentów. Konstruktor 61 unikalnych konstrukcji urządzeń mechatronicznych. Autor licznych wdrożeń przemysłowych.



e-mail: [bartys@mchtr.pw.edu.pl](mailto:bartys@mchtr.pw.edu.pl)