



POLITECHNIKA OPOLSKA

**KATEDRA MECHANIKI I PODSTAW
KONSTRUKCJI MASZYN**

Instrukcja do ćwiczeń laboratoryjnych
z elementów analizy obrazów

**Zastosowanie Vision Assistant
do selekcji elementów z zastosowaniem techniki OCR.**

Opracował: dr inż. Roland Pawliczek

Opole 2016

Publikacja na prawach rękopisu

1. Cel ćwiczenia

Celem ćwiczenia jest zapoznanie się z procedurą wykorzystania techniki OCR do identyfikacji obiektów opisanych tekstowo z zastosowaniem modułu Vision Assistant.

2. Informacje wstępne.

OCR (ang. Optical Character Recognition) to zestaw technik lub oprogramowanie służące do rozpoznawania znaków i całych tekstów w pliku graficznym o postaci rastrowej. Zadaniem OCR jest zwykle rozpoznanie tekstu w analizowanym obrazie. W najbardziej zaawansowanej postaci są to systemy, które umożliwiają przenoszenie dokumentów dostępnych jako obraz (zwykle skanowany) do oprogramowania służącego edycji z uwzględnieniem formatowania tekstu, tabel itp.

W praktyce przemysłowej informacje tekstowe z etykiet, nadruków na obiektach zwykle służą ich identyfikacji i przeprowadzaniu np. selekcji czy kontroli.

W algorytmach OCR program przetwarzający dzieli stronę na poszczególne elementy, identyfikuje miejsca, w których znajduje się tekst. W dalszej kolejności tekst dzielony jest na coraz mniejsze elementy: akapity, wersy, słowa i znaki. Dokument zapamiętywany jest jako zbiór pojedynczych znaków. Ważnym elementem techniki OCR jest wykorzystanie wzorców. Wymaga to poprzedniego przygotowania baz danych.

Technika rozpoznawania znaków z obrazów rastrowych wymaga odpowiedniego przygotowania materiału do obróbki. Znaczenie ma tutaj rozdzielczość, kontrast → jakość obrazu do analizy. Przygotowując materiały wyjściowe do rozpoznania tekstu (np. formę etykiet) skuteczność rozpoznania mogą podnieść:

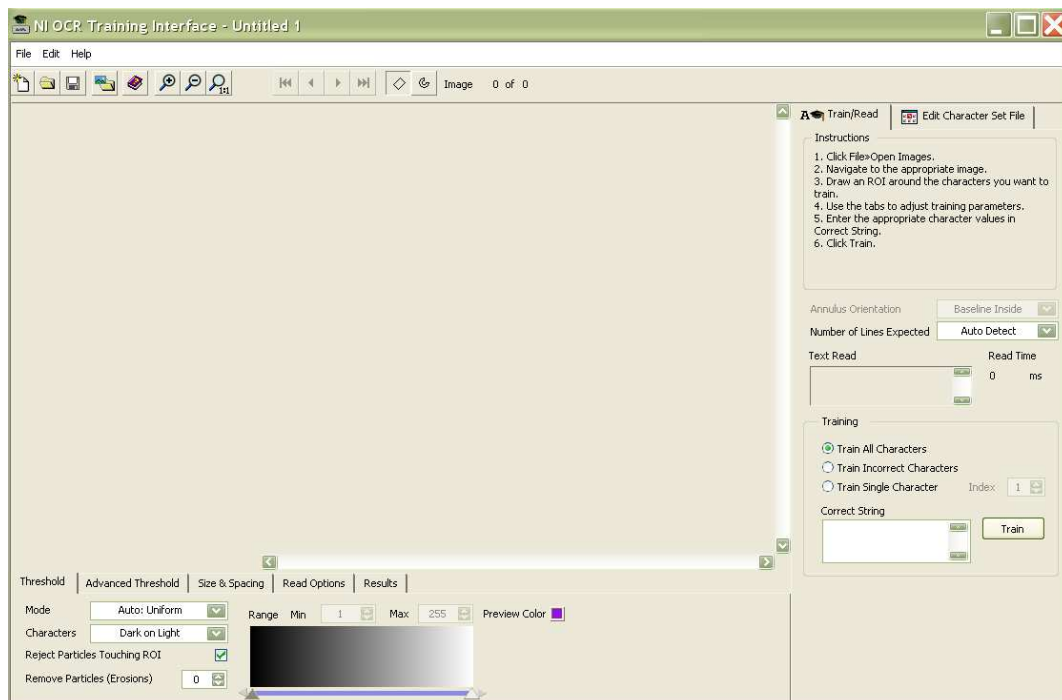
- tło dla tekstu powinno być jasne, bez defektów (zagniecenia, zabrudzenia),
- dokument należy skanować z rozdzielczością co najmniej 300 dpi lub wyższą
- obraz przeznaczony do OCR powinien być zapisany w odcieniach szarości, najlepiej w formacie nieskompresowanym (TIFF, GIF, BMP, PNG)
- należy przetestować ustawieniom jasności i kontrastu w programie skanującym, ewentualnie wykonać wstępną obróbkę obrazu przed zastosowaniem procedur OCR,
- najtrudniej interpretuje się drobne znaki np. kropki, przecinki, lub litery podobne (j-i, n-m, b-h).

3. Tworzenie bazy – wzorca znaków.

Tworzenie bazy znaków dla funkcji OCR możliwe jest za pomocą dedykowanego narzędzia dostępnego z poziomu systemu operacyjnego:

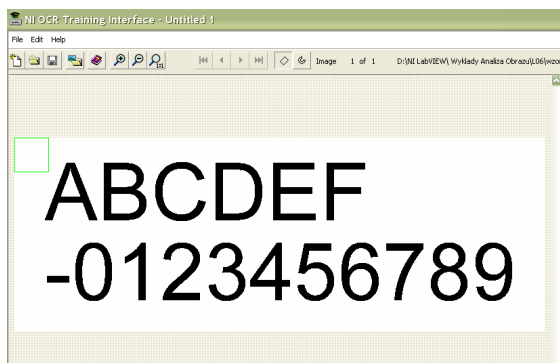
START/Wszystkie programy/National Instruments/Vision/Utilities/Ocr Trainig.

Uruchomiony zostanie program, którego główne okno przedstawione jest poniżej.

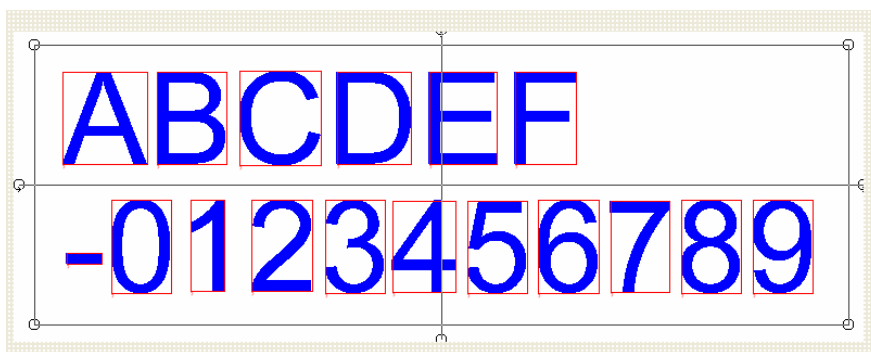


Proces tworzenia bazy znaków odbywa się na wcześniej przygotowanym zestawie znaków zawierającym wszystkie, niezbędne do poprawnej identyfikacji litery, cyfry i znaki specjalne, przy czym obraz wzorcowy powinien być przygotowany w formacie i przy parametrach jak system skanujący.

W tym przypadku jest to plik graficzny **wzorzec.png** wczytany za pomocą opcji *File/Open Images...* Na ekranie pojawi się zestaw znaków do nauki.



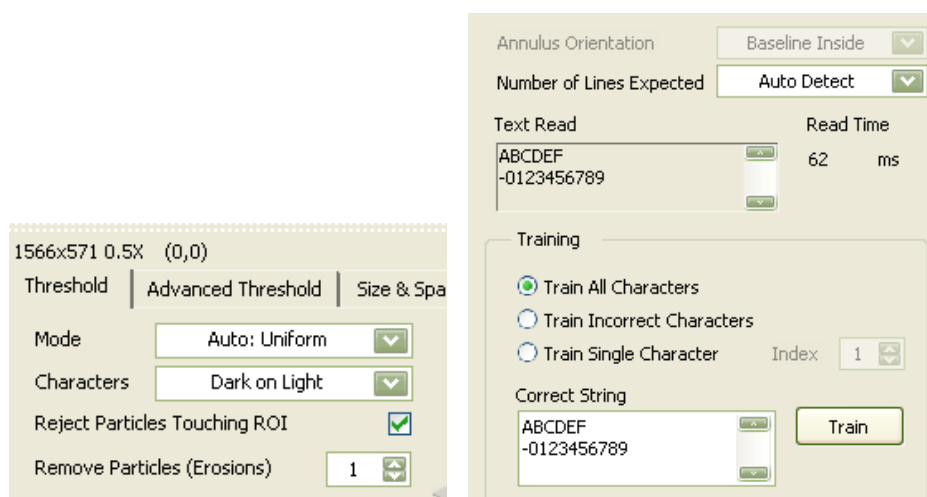
Następnie należy zaznaczyć obszar obejmujący wszystkie znaki, które powinny zostać odseparowane przez program.



W zakładce *Threshold* należy ustawić pola: *Mode=Auto:Uniform*, *Characters=Dark on Light*.

Opcja *Reject Particles Touching ROI* pozwoli odrzucić wszystkie obiekty przypadkowo objęte linią selekcji obszaru znaków do nauki. Opcja *Remove particles (Erosions)=1* wykona proces filtracji nazywany erozją, który pozwoli lepiej odseparować krawędzie znaków od tła.

Proces tworzenia bazy polega na wpisaniu w pole *Correct String* poprawnego ciągu znaków tworzący wzorec. Aby zapamiętać wszystkie znaki należy wybrać opcję *Train All Characters* a następnie nacisnąć klawisz *Train*.



Ciąg znaków rozpoznany przez system zostanie wyświetlony w oknie *Text Read*. Można tu sprawdzić poprawność tworzenia bazy danych.

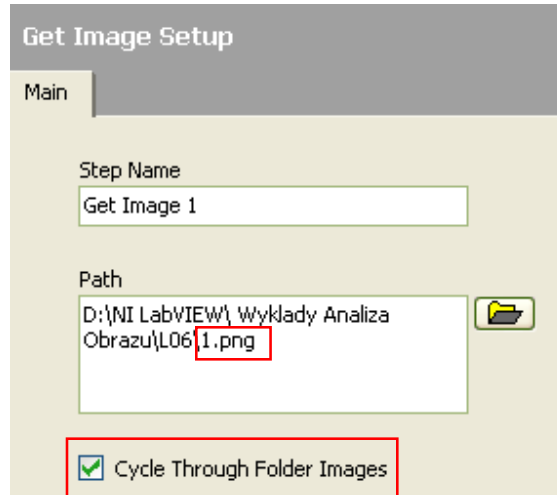
UWAGA: utworzoną bazę zapisać za pomocą opcji *File/Save Character Set File* → zostanie utworzony plik z rozszerzeniem „.abc”.

4. Skrypt przetwarzania obrazów.

4.1 Wczytanie obrazu do analizy.

Uruchomić moduł **Vision Assistant**. Z palety funkcji *Processing Function: Image* wybrać funkcję **Get Image**.

W oknie konfiguracyjnym funkcji odszukać plik **1.png** i zaznaczyć opcję *Cycle Thru Folder Image*. Zatwierdzić wybór.



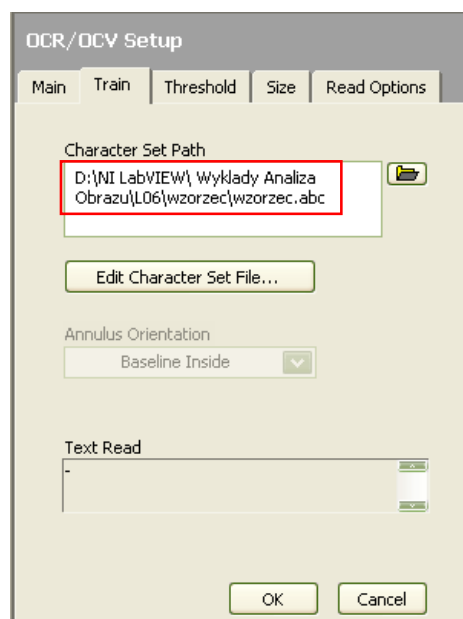
4.2 Konfiguracja OCR.

Z palety funkcji *Processing Function: Identification* wybrać opcję *OCR/OCV: Reads characters in a region of the image*.



OCR/OCV: Reads characters in a region of the image.

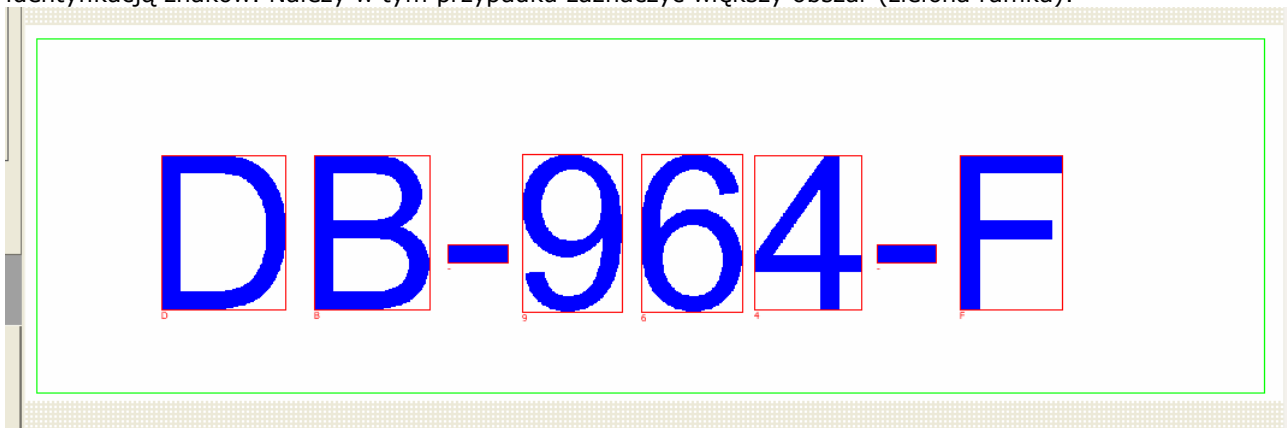
W oknie konfiguracyjnym w zakładce *Train* należy wczytać utworzony wcześniej plik z utworzoną bazą danych znaków (w tym przypadku **wzorzec.abc**). Klawisz *Edit Character Set File...* pozwala edytować i korygować wzorzec ze znakami uruchamiając wcześniej używany program do tworzenia baz danych.



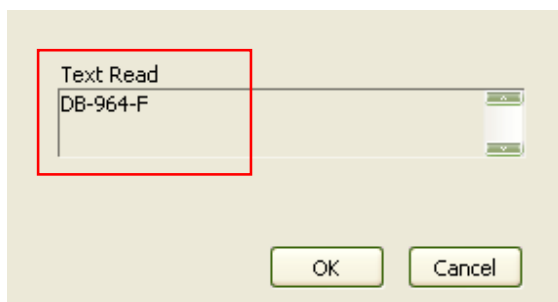
Uwaga: pozostałe zakładki pozwalają zmieniać parametry rozpoznawania. Stosowane są w przypadku, gdy pojawiają się obrazy w sposób znaczący odbiegające od przyjętego standardu i wyniki rozpoznania są błędne.

Następnie należy zaznaczyć obszar, który zawierać będzie sekwencję znaków do rozpoznania.

Uwaga: jeżeli na kolejnych obrazach pozycja znaków nie będzie taka sama mogą pojawić się problemy z identyfikacją znaków. Należy w tym przypadku zaznaczyć większy obszar (zielona ramka):



Znaki zostaną rozpoznane i pokazane w polu *Text Read* i dodatkowo w tabeli zostaną zidentyfikowane wszystkie znaki wraz z opisem położenia i wartości współczynnika *Threshold*, który określa stopień zgodności rozpoznanego znaku ze wzorcem (im bliżej 1000 tym lepiej).



Results ...	D	B	-	9	6	4	-	F
Classification Score	1000	1000	1000	964	1000	999	1000	988
Identification Score								
Left	149	317	465	547	677	802	937	1028
Top	143	143	241	142	142	143	241	143
Width	138	127	65	109	111	117	65	114

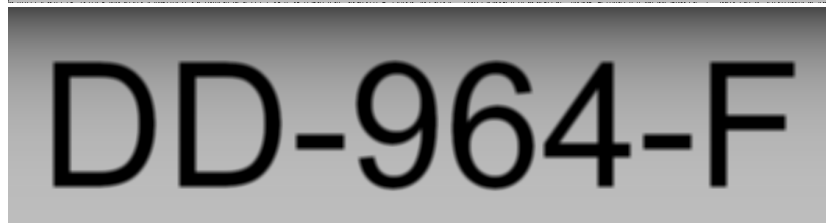
Wykonać identyfikację kodu dla wszystkich obrazów **1.png – 4.png**.

5. Zadanie.

Dodatkowe pliki obrazów **a1.png – a4.png** zawierają kody ze zniekształceniami:



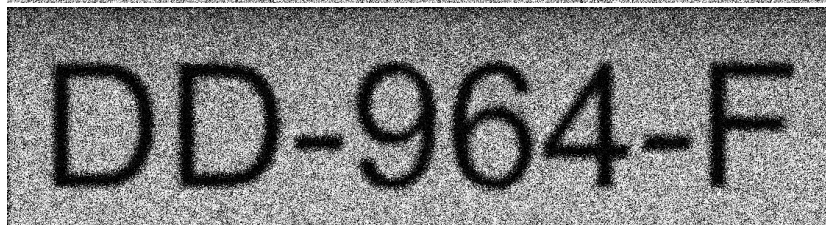
Erozja krawędzi znaków.



Rozmycie krawędzi znaków.



Wyraźne znaki z zaszumionym tłem.



Erozja krawędzi z zaszumionym tłem.

Wykorzystując utworzony skrypt przetwarzania obrazu sprawdzić, które zniekształcenia wpływają na rezultaty odczytu kodu. Czy możliwa jest korekta obrazu tak, aby ewentualne zniekształcenia usunąć i przeprowadzić poprawny odczyt kodu?

Wskazówka: przed procedurą OCR sprawdzić działanie funkcji z palety Grayscale: *Lookup Table, Filters, Gray Morphology*.

Obserwacje i przetworzone obrazy opisać i zamieścić w raporcie.