

Laboratorium Rozpoznawania Obrazów – Ćwiczenie #3 & #4

Rozpoznawanie ręcznego pisma blokowego

Termin oddawania: **16.04.2012, 12.04.2012, 23.04.2012, 19.04.2012**

W tym ćwiczeniu Państwa zadaniem będzie rozpoznawanie cyfr pisanych ręcznie. Do pierwszej fazy zadania – przygotowania klasyfikatora liniowego – wykorzystamy dane z bazy danych MNIST (<http://yann.lecun.com/exdb/mnist>). Zbiory uczący i testowy można pobrać spod podanego wyżej adresu. Dodatkowo są dostępne w laboratorium 313 oraz na serwerze Galera.

Pracę klasyfikatora „u klienta” będzie można przećwiczyć korzystając z drugiego zbioru testowego, zawierającego cyfry pisane przez studentów WEiTl. Ten zbiór, także w formacie MNIST, jest dostępny w laboratorium 313 oraz na serwerze Galera.

Warto wiedzieć, że obrazy są normalizowane po zeskanowaniu w następujący sposób:

1. Prostokąt zawierający czarno-biały obraz znaku zeskanowanego w rozdzielczości 300 dpi jest skalowany proporcjonalnie do prostokąta o większym z wymiarów równym 20. W trakcie skalowania obraz jest zamieniany na skalę szarości (proporcjonalnie do liczby oryginalnych pikseli pierwszego planu przypadających na jeden piksel obrazu po przeskalowaniu).
2. Jest wyznaczany środek ciężkości przeskalowanego znaku, a znak jest umieszczany w obrazie 28x28 pikseli tak, żeby środek ciężkości znalazł się na środku tego większego obrazu.

Drugi zestaw znaków został poddany dodatkowo usuwaniu przekosu.

Łączy do plików na serwerze Galera są następujące:

<http://galera.ii.pw.edu.pl/~rkz/mnist.zip> (oryginalne dane MNIST)

<http://galera.ii.pw.edu.pl/~rkz/testdigits.zip> („rzeczywisty” zbiór testowy w formacie MNIST)

<http://galera.ii.pw.edu.pl/~rkz/reference.zip> (rozwiązanie referencyjne - warto w nim sprawdzić, jak odczytywać dane – lub po prostu przejąć z niego funkcję odczytu danych)

Państwa zadaniem, jest przygotowanie klasyfikatora korzystającego z **klasyfikatorów liniowych** rozróżniających konkretne cyfry. Oprócz jakości klasyfikacji na zbiorach testowych (zbiorze pochodzącym z bazy danych MNIST i „rzeczywistym” zbiorze testowym) należy podać macierze pomyłek. Dodatkowe zadanie polega na zaproponowaniu sposobu wyznaczania „ufności” do decyzji klasyfikatora (chodzi o wartość z przedziału 0..1 opisującą pewność, że decyzja klasyfikatora jest poprawna).

Rozwiązaniem referencyjnym, jest klasyfikator najbliższego sąsiedztwa (w wersji 5-NN z głosowaniem absolutną większością) działający na cechach otrzymanych bezpośrednio z obrazów znaków metodą składowych głównych (PCA). W testach użyłem 40 składowych głównych. Wyniki klasyfikacji podsumowuje poniższa tabela (choć ciekawy wgląd w klasyfikację może dać analiza macierzy pomyłek).

	Zbiór testowy MNIST			„Rzeczywisty” zbiór testowy		
	OK.	Błąd	Odrzucenie	OK.	Błąd	Odrzucenie
Jakość klasyfikacji *)	97.21%	2.02%	0.77%			

*) Wyniki z 23 marca – mogą ulec zmianie ☺

Zadanie można podzielić na kilka części:

1. Przygotowanie podstawowego algorytmu wyznaczającego parametry płaszczyzny decyzyjnej dla zadanego zbioru uczącego, zawierającego dwie klasy. Ten wstępny krok najlepiej przeprowadzić na zbiorze dwuwymiarowym, który będzie łatwo zwizualizować i pokazać wyznaczone rozwiązanie. Oprócz przejrzania materiałów z wykładu wartościowe będzie zajrzenie do Numerical Recipes in C (rozdział 10, „Minimization and Maximization of Functions”).
2. Testy algorytmu dla wielowymiarowych danych cyfr. Przeprowadzenie tego testu ma na celu sprawdzenie wydajności algorytmu, oraz jego zachowania w przestrzeni wielowymiarowej. Dane źródłowe mogą być użyte bezpośrednio (tzn. każdy piksel oryginalnych obrazów jest traktowany jako cecha), lub po redukcji wymiarowości algorytmem PCA (.).
3. I tu pojawia się najciekawszy moim zdaniem punkt, bo klasyfikator liniowy rozróżnia między dwoma klasami, a cyfr jest 10. W jaki sposób złożyć wyniki pojedynczych klasyfikatorów? Ile tych pojedynczych klasyfikatorów powinno być? Czy klasyfikator złożony będzie mógł się uchylić od decyzji? (Czyli zamiast wskazania konkretnej klasy odpowiedzieć „nie wiem”). Jeśli tak, to w jakich przypadkach? Wbrew numeracji, nad tym punktem warto zastanowić się najpierw, bo może mieć wpływ na to, jakie klasyfikatory będą w ogóle potrzebne.
4. Ostatni krok – którym wg mnie warto zająć się kiedy klasyfikator będzie gotowy – jest zdefiniowanie sposobu wyznaczania ufności dla decyzji klasyfikatora. Na koniec należy sprawdzić, na ile poziom ufności jest skorelowany z poprawnymi/błędnymi decyzjami klasyfikatora.

W sprawozdaniu proszę zamieścić:

1. Opis metody klasyfikacji znaków przy użyciu klasyfikatorów liniowych.
2. Opis algorytmu wyznaczania parametrów płaszczyzny decyzyjnej.
3. Dane dotyczące jakości klasyfikacji każdego z wykorzystywanych klasyfikatorów liniowych.
4. Dane dotyczące jakości klasyfikacji cyfr, z wnioskami wynikającymi z analizy **macierzy pomyłek**.
5. Sposób liczenia poziomu ufności i wyniki badania korelacji tego współczynnika z decyzjami klasyfikatora.

Uwaga: Pasjonaci mogą powtórzyć procedurę (przy ustalonych parametrach klasyfikatora) biorąc jako zbiór uczący całą bazę MNIST (70000 cyfr) i zastosować ją do klasyfikacji drugiego zbioru testowego.