

POPULARNONAUKOWE STRESZCZENIE PROJEKTU

W ostatniej dekadzie obserwujemy powszechniejszą obecność oprogramowania do automatycznego wykrywania lub śledzenia obiektów na obrazach cyfrowych czy też sekwencjach wideo, np.: detektory twarzy (w aparatach telefonów komórkowych, galeriach portalu Facebook), detektory osób i pojazdów (w systemach monitoringu), detektory znaków drogowych (w nowoczesnych samochodach). W dużej większości tych systemów ważnym elementem składowym są tzw. *cechy Haara* zaproponowane w pracach (Viola i Jones, 2001; Viola i Jones, 2004). Są to proste cechy różnicowe, które służą do opisu analizowanego okna obrazu i mogą być interpretowane jako zgrubne kontury — np.: różnice pomiędzy średnią jasnością linii czoła i oczu, nosa i policzków, itp. Odpowiednio duży zestaw takich prymitywnych cech (np. rzędu 10^4) może stanowić wystarczającą podstawę do odróżnienia pożądaných obiektów od tła czy też od innych obiektów.

Z algorytmicznego punktu widzenia istotną zaletą cech Haara jest możliwość ich bardzo szybkiego wyznaczenia po przygotowaniu tzw. *obrazu całkowego*. Jest to pomocnicza tablica przechowująca skumulowaną funkcję jasności obrazu i to właśnie ona jest tak naprawdę sednem pomysłu. Wówczas, niezależnie od tego czy badane okno jest rozmiaru np. 24×24 pikseli czy też 256×256 , liczba operacji potrzebna do wyznaczenia pojedynczej cechy pozostaje taka sama, co informatycy nazywają *stałą* złożonością obliczeniową — $O(1)$. Ten fakt jest ważny, ponieważ gęsta procedura detekcyjna skanująca obraz oknem przesuwającym wymaga wykonania obliczeń dla bardzo wielu okien. Zależy to od rozdzielczości obrazu i pewnych nastaw programistycznych (liczba skal, skoki okna detekcyjnego, itp.), niemniej jednak w typowo spotykanych zastosowaniach liczba badanych okien dla pojedynczego obrazu jest rzędu 10^5 lub nawet 10^6 . W konsekwencji analiza pojedynczego okna musi odbywać się w czasie rzędu mikrosekund lub krótszym, tak aby cała procedura zakończyła się w rozsądnym czasie.

Pomimo szybkości, detektory oparte na cechach Haara nie są wolne od błędów. Można zaobserwować sporadyczne braki wskazań na właściwych obiektach lub fałszywe wskazania pozytywne. Istnieją nieliczne urozmaicenia pomysłu bazowego (np. deskryptor HOG oparty na gradientach kierunkowych czy też cechy związane z wariancją / kowariancją szarości) będące technikami również o złożoności stałej, ale i one pozostawiają pewne pole do poprawy dokładności ze względu na ogólną prostotę wspomnianych cech (niską informacyjność). Powyższe powody były argumentem za podjęciem przez autorów tej tematyki badawczej.

Głównym celem naukowym projektu jest: *opracowanie nowych całkowitych reprezentacji obrazów i bazujących na nich algorytmów pozwalających na ekstrakcję cech (momentów Fouriera, momentów Fouriera-Mellina, momentów statystycznych) z okien obrazu w czasie stałym — $O(1)$ — na potrzeby zadań uczenia maszynowego i detekcji*. A zatem ambicją autorów jest poszukiwanie możliwości ekstrakcji bardziej zaawansowanych cech (o lepszych własnościach aproksymacyjnych), ale nadal z wykorzystaniem szybkiego podejścia bazującego na stałej złożoności obliczeniowej.

Poza pracami matematycznymi i algorytmicznymi w projekcie przeprowadzane będą badania z zakresu uczenia maszynowego obejmujące: naukę detektorów na dużych zbiorach danych poprzez algorytmy boostowane, eksperymenty na różnych wariantach tzw. słabych klasyfikatorów, wykorzystanie technik przyspieszających („weight trimming”) i regularyzacyjnych. Otrzymane rezultaty będą oceniane dwójako. Po pierwsze poprzez miary dokładności klasyfikatorów (m.in. czułość, FAR, F_1 -score, AUC), po drugie poprzez jakość rekonstrukcji obrazów uzyskiwanych za pomocą cech. Rekonstrukcje to inaczej przybliżenia obrazów lub okien obrazu, które można zrealizować poprzez rozwinięcia w sumy częściowe szeregu momentów. O ile rekonstrukcje nie są potrzebne bezpośrednio w samym zadaniu detekcji, to mogą być one dobrym miernikiem ekstrahowanych cech pod względem ich przydatności właśnie do detekcji.