

BOGUSŁAW CYGANEK

Metody i algorytmy rozpoznawania obiektów w obrazach cyfrowych

Streszczenie

Rozpoznawanie obiektów należy do podstawowych zadań dziedziny widzenia komputerowego, którego głównym celem jest nauczenie komputerów obserwacji oraz rozumienia zawartości scen w sposób zbliżony, a czasami nawet przewyższający zdolności człowieka. Komputerowe metody rozpoznawania obiektów znajdują coraz szersze zastosowanie w takich dziedzinach jak rozpoznawanie twarzy, systemy obserwacji i nadzoru, przetwarzanie obrazów medycznych, multimedia, automatyczne prowadzenie pojazdów, analiza dokumentów, rozpoznawanie rysunków odręcznych oraz wielu innych.

Monografia prezentuje opracowane przez autora metody oraz algorytmy rozpoznawania obiektów w obrazach cyfrowych. Są one uzupełnione opisem ich podstaw teoretycznych, jak również odnośnikami do innych prac w tej dziedzinie.

Monografia jest podzielona na sześć rozdziałów oraz dodatki, z których pierwszy stanowi wprowadzenie w prezentowaną tematykę. Rozdział 2 zawiera przegląd różnorodnych metod tensorowych przydatnych w dziedzinie widzenia komputerowego. Często stanowią one trzon opracowanych przez autora innych metod rozpoznawania obiektów. Specjalną uwagę poświęcono tensorowi strukturalnemu, jak również jego wersjom zależnym od wymiarowości sygnału wejściowego. Dokładnie przedstawione zostały również zagadnienia implementacyjne tego tensora.

W wielu zastosowaniach dane tensorowe wymagają wstępnej filtracji. Jest to zagadnienie trudne ze względu na wielowymiarowość tego typu danych. W wyniku poszukiwania rozwiązania tego problemu została opracowana oryginalna metoda filtracji tensorów z wykorzystaniem filtrów bazujących na statystykach porządkowych. Umożliwia ona efektywną filtrację tensorów o dowolnym wymiarze.

W rozdziale 3 zaprezentowane zostały statystyczne metody klasyfikacji obiektów. Specjalną uwagę poświęcono grupie tzw. metod "miękkich obliczeń". Należą do nich sztuczne sieci neuronowe, klasyfikatory z wektorami wspierającymi, jak również logika rozmyta. Te bardzo efektywne, a jednocześnie wysoce intuicyjne techniki bardzo często są stosowane w innych metodach opracowanych przez autora. Ostatnią z diskutowanych w tym rozdziale jest metoda przesunięcia średniego. Ta bardzo efektywna i ważna technika, również wywodząca się ze statystyki matematycznej, znajduje zastosowanie głównie w segmentacji obrazów oraz do śledzenia obiektów w obrazach.

Zagadnienia detekcji obiektów oraz ich śledzenia dyskutowane są w rozdziale 4. Zaczyna się on prezentacją dwóch oryginalnych rozwiązań bezpośredniej klasyfikacji pikseli. Pierwsze z nich, które bazuje na wnioskowaniu rozmytym, zwiększa stopień detekcji obrazów ludzkiej skóry i w konsekwencji – działając jako filtr wstępny – zwiększa skuteczność metod detekcji ludzkich twarzy. Druga z metod umożliwia bezpośrednią segmentację pikseli należących do znaków drogowych. Ma ona dwa warianty, pierwszy bazujący na logice

rozmytej, drugi na klasyfikatorach z wektorami wspierającymi. Ten ostatni pozwala otrzymać najdokładniejsze wyniki, kosztem jednak bardziej złożonej implementacji oraz nieco dłuższego czasu działania. Następnie diskutowana jest metoda detekcji podstawowych kształtów z wykorzystaniem tensora strukturalnego. Innowacyjna wersja tej metody, zwana metodą „zapisu w górę”, pozwala na szybką detekcję krzywych wypukłych na podstawie tensora strukturalnego. Jest ona konkurencyjna do szeroko stosowanej transformaty Hougha. W dalszej części przedstawione zostały nowatorskie metody detekcji figur. Umożliwiają one detekcję trójkątów, prostokątów, rąbów, itp. na podstawie detekcji ich narożników. Detekcja innych kształtów, np. elips, jest możliwa z wykorzystaniem opracowanej przez autora metody adaptacyjnie rosnącego okna. Działanie wszystkich z wymienionych metod zostało zweryfikowane w systemie detekcji znaków drogowych.

W końcu rozdziału 4 zaprezentowane zostały dwa systemy śledzenia obiektów. Pierwszy z nich umożliwia śledzenie znaków drogowych z wykorzystaniem rozmytej wersji algorytmu przesunięcia średniego. Druga jest wszechstronną platformą śledzenia dowolnych obiektów w filmach, bazującą na nowatorskim połączeniu koloru oraz cech strukturalnych tych obiektów.

Rozpoznawanie obiektów, które jest swoistą syntezą metod zaprezentowanych w poprzednich rozdziałach, diskutowane jest w rozdziale 5. Rozpoczyna się on prezentacją nowatorskiej metody rozpoznawania obiektów sztywnych na podstawie histogramów fazowych tensora strukturalnego, obliczonych w morfologicznej przestrzeni skal. Metoda ta jest niezmiennicza ze względu na transformację afiniczną. Następnie diskutowana jest metoda korelacji szablonów typu log-polar w gaussowskiej przestrzeni skal. W dalszej części zaprezentowano metody bazujące na sztucznych sieciach neuronowych. Pierwsza w tej grupie jest metoda wykorzystująca probabilistyczną sieć neuronową operującą na niezmiennikach momentów afinicznych. Druga z nich stanowi nowatorską konstrukcję współpracujących eksperckich klasyfikatorów z sieciami Hamminga operujących w grupie deformowalnych modeli. Cały zespół tych klasyfikatorów jest zarządzany poprzez moduł arbitrażowy pracujący w trybie wyłaniania jednego zwycięskiego neuronu, uzupełniony dodatkowo o nowy mechanizm wspierania grupy jednomyślnych ekspertów. Taka hybrydowa konstrukcja wykazała dużą dokładność, jak również znaczną wszechstronność, skutecznie pracując w kompletnym systemie rozpoznawania polskich znaków drogowych, których zdefiniowanych jest ponad dwieście różnych wzorów. Na zakończenie tego rozdziału diskutowane są metody konstrukcji oraz rozwoju kompletnych systemów widzenia komputerowego.

Ostatni rozdział niniejszej monografii stanowią dodatki, w których przedstawione zostały podstawy morfologicznych przestrzeni skal, morfologicznych operatorów tensorowych, jak również geometryczne właściwości form kwadratowych. Monografię kończy spis literatury oraz skorowidz haseł.

BOGUSŁAW CYGANEK

Methods and Algorithms of Object Recognition in Digital Images

Summary

Object recognition in digital images is a key problem of Computer Vision (CV) which aims in making computers to perceive and interpret space, based on visual information, in a way that is close, or even excels, human abilities. Recognition methods find broad and still expanding application in such areas as face recognition, surveillance systems, medical imaging, multimedia, vehicle driving, document analysis, hand drawings recognition, and many more.

The monograph presents methods and algorithms of object recognition in digital images developed by the author. These are supplemented with their theoretical foundations, as well as with references to other works in this area.

The monograph is divided into six chapters and an appendix. After the introduction, in Chapter 2 an overview is presented of diverse tensor methods suitable to the tasks of CV. These are used frequently in the recognition methods developed by the author. Special attention was devoted to the structural tensor, as well as to its different versions depending on dimensionality of the input signal. Its implementation issues are also discussed.

In many applications tensor data prior to processing need to be filtered. This makes some difficulties due to high dimensionality of tensor fields. For this purpose an original method of order statistic filter was developed by the author. It allows efficient filtering of any dimensional tensors.

In the following Chapter 3 an insight into the statistical framework for object classification is provided. Special attention was devoted to the group of soft methods. These are different neural networks, support vector machines, as well as fuzzy logic. These very powerful and intuitive tools are frequently used in the methods developed by the author. The last discussed in this chapter is the mean shift method. This very important and versatile technique, also built upon the statistical framework, finds application mostly in image segmentation and object tracking.

Object detection and tracking are discussed in Chapter 4. It starts with presentation of the two original solutions to the direct pixel classification. The first one, which is based on the fuzzy reasoning, enhances human skin detection and, in consequence, can be used to facilitate face detection. The second method allows direct segmentation of pixels belonging to the road signs. This has two variants, one based on fuzzy logic, the second on the support vector machines. The latter shows superior results at a cost of more complicated implementation and longer run time. Then the method of detection of basic shapes from the structural tensor is discussed. An innovative author's version allows fast detection of curves from the tensor in the up-write mode. This method is competitive to the Hough transform. Novel methods of figure detection are discussed next. They allow detection of triangles, rectangles, diamond-like shapes, etc. based on detection of their salient points. Detection of other

shapes, such as an ellipse, can be done with the new adaptive window growing method. All these were verified in the system of road signs detection.

At the end of Chapter 4 two systems of object tracking are discussed. The first one allows tracking of the road signs with the fuzzy version of the mean shift method. The second is a general framework for object tracking in films based on a novel way of composition of color and structural features.

Object recognition, which constitutes a synthesis of the methods from the previous chapters, is dealt with in Chapter 5. It starts with detailed description of the novel method of solid object recognition from phase histograms of the structural tensor, computed in the morphological scale-space. The method is invariant to affine transformations. Next, an original method of the log-polar template matching in the Gaussian scale-space is presented. Then the methods operating with the neural classifiers are discussed. The first one relies on the original connection of the affine moment invariants classified with the probabilistic neural network. The second presents an innovative application of an ensemble of Hamming neural expert-classifiers operating within the group of deformable models. The whole ensemble is controlled by the arbitration unit operating in the winner-takes-all fashion, endowed with the novel mechanism of support within the group of unanimous experts. This hybrid connection showed to be very accurate and versatile in the complete system of recognition of the Polish road signs, which set contains more than two hundred of different patterns. Finally, methods of construction and development of the CV systems are discussed.

The last chapter of this monograph provides additional theoretical background on morphological scale-space, morphological tensor operators, and geometry of the quadratic forms. The monograph ends with the bibliography and index.