

Streszczenie

Podział na bloki w kodowaniu wizji z wykorzystaniem sztucznych sieci neuronowych

W niniejszej rozprawie przedstawiono wyniki badań w dziedzinie kodowania wideo, gdzie obrazy tworzące sekwencję wizyjną są poddawane kompresji, co wykorzystywane jest w systemach telewizyjnych, platformach VoD (Video-on-Demand) udostępniających filmy na żądanie i podobnych zastosowaniach. Badania przedstawione w tej rozprawie skupiają się na rozwoju algorytmów sterowania koderem. Szczególny nacisk położono na wyznaczanie podziałów bloku CTU (Coding Tree Unit), co jest najbardziej złożonym obliczeniowo etapem kodowania. Celem badań jest opracowanie algorytmu wyznaczania podziałów bloku CTU, który znacznie zmniejszy złożoność obliczeniową kodera, przy jednoczesnym zachowaniu wydajności kompresji w odniesieniu do rozwiązania zastosowanego w oprogramowaniu referencyjnym. W związku z tym zbadano możliwość wykorzystania sztucznych sieci neuronowych (ANN) w algorytmie podziału bloku CTU.

Na początku rozprawy przedstawiono opis procesu podziału bloków CTU w technice HEVC. W opisie uwzględniono kontekstowość w podejmowaniu decyzji. Przedstawiony został przegląd istniejących algorytmów wyznaczania podziałów bloku CTU, wskazując aspekty, które wymagają ulepszeń. Równocześnie dokonano przeglądu metod pozwalających na sterowanie kompromisem pomiędzy czasem kodowania i wydajnością kodowania, gdzie wskazano na skomplikowanie sterowania tym kompromisem w istniejących rozwiązaniach.

W ramach badań rozważano dwa podejścia do problemu podziału bloków CTU: jeden rozważający decyzję o podziale na poziomie bloków Coding Unit (CU), oraz drugi rozpatrujący decyzje na poziomie zarówno bloków Coding Unit (CU) jak i Prediction Unit – PU). Proponowane sieci neuronowe zaprojektowano do łącznego estymowania prawdopodobieństw głębokości podziału dla poszczególnych obszarów w CTU. Do wytrenowania takiej sztucznej sieci neuronowej przygotowano dedykowany zbiór danych uczących. W rozprawie opisano proces tworzenia architektury sztucznych sieci neuronowych, wraz ze szczegółową analizą wyników trenowania i wydajności ostatecznych modeli.

Następnie, w niniejszej rozprawie, przedstawiono autorskie algorytmy decyzyjne, wykorzystujące prawdopodobieństwa estymowane przy pomocy sztucznych sieci neuronowych. Algorytmy te zdefiniowano w dwóch wariantach: twardo decyzyjnym i miękko decyzyjnym. W rozprawie wykazano, że jeden z proponowanych algorytmów pozwala na prostą kontrolę kompromisu między czasem kodowania a wydajnością kodowania używając tylko jednego parametru.

Analiza porównawcza z najnowszymi rozwiązaniami wykazała, że proponowane algorytmy partycjonowania oferują najlepszy kompromis między skróceniem czasu kodowania a wydajnością kodowania.

Biorąc pod uwagę kontrolę nad kompromisem między czasem kodowania a wydajnością kodowania, spośród metod znanych w literaturze, zaproponowana metoda zapewnia najlepsze wyniki i najprostszą kontrolę nad tym kompromisem. Dodatkowo wprowadzono autorską metrykę do szybkiego porównywania takich metod, której wyniki pokrywają się z innymi, dobrze znanymi metodami.

Na końcu rozprawy, zbadano wpływ efektów kontekstowych na decyzje dotyczące podziału bloku CTU. Kluczowe osiągnięcia w tym obszarze to:

- Zaproponowany autorski eksperyment do określania wpływu kontekstu kodowania na decyzje dotyczące podziału bloku CTU.
- Modyfikacje zaproponowanej sztucznej sieci neuronowej do przetwarzania informacji kontekstowych.

- Opracowana przez autora metoda trenowania sieci neuronowej poprzez rozszerzenie etykiet uczących (ground truth augmentation).
- Autorska metoda globalnej optymalizacji decyzji o podziałach bloków CTU, nieistotnie zwiększająca czas kodowania.

Wszystkie eksperymenty przedstawione w niniejszej rozprawie zostały przeprowadzone przy użyciu opracowanej przez autora zmodyfikowanej wersji oprogramowania modelowego dla techniki HEVC. Modyfikacja ta, umożliwiającą szybką implementację algorytmów podziału bloków CTU opartych na sztucznych sieciach neuronowych, została udostępniona na zasadach otwartego dostępu (open-access).