

Michał Podpora

Wydział Elektrotechniki, Automatyki i Informatyki
Instytut Automatyki i Informatyki

WYKORZYSTANIE DYSKRETNEJ TRANSFORMATY FALKOWEJ PRZY KOMPRESJI OBRAZU SYSTEMU WIZYJNEGO ROBOTA MOBILNEGO

Abstract: In this article author considers using DWT transform and RLE compression in a specific machine vision approach. It might be especially useful in HTM-based processing of robot's vision system data.

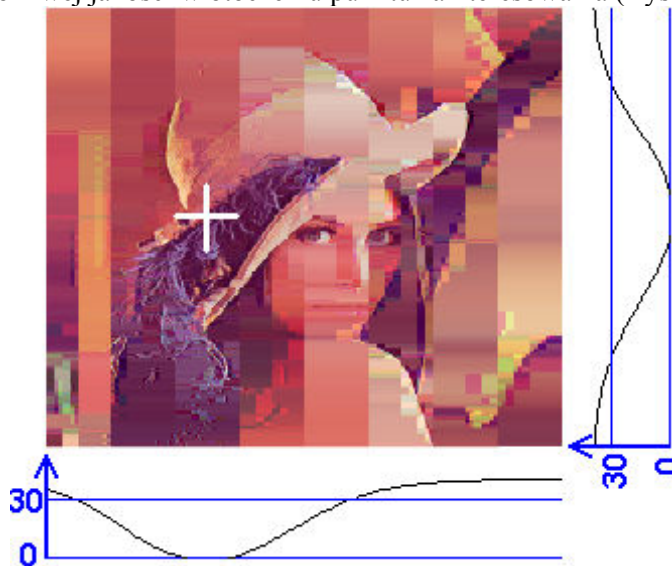
Pomysł zastosowania klastra komputerowego do zwiększenia możliwości robota mobilnego nie jest pomysłem nowym. W historii odnotowano już kilka prób zbudowania takiego systemu. Niestety sieć łącząca węzły klastra okazywała się być przeciążona przesyłanymi danymi wizualnymi. Wszelkie próby wprowadzenia kompresji obrazu powodowały pogorszenie jakości, co z kolei utrudniało jego analizę w węzłach.

Istnieje jednak rozwiązanie tego problemu. Tak jak ludzkie oko nie przesyła do mózgu obrazu w postaci rastrowej, tak samo system wizyjny mógłby być zaprojektowany tak, aby umożliwiał analizę obserwowanej sceny „detal po detal”. Podejście fragmentarycznego analizowania sceny jest od wielu lat znane naukowcom (funkcjonuje pod nazwą „active vision”), ale zazwyczaj rozumiane jedynie jako pomniejszenie obszaru zainteresowania do mniejszego prostokątnego fragmentu całej bitmapy [1].

Redukcja rozmiaru przesyłanej porcji danych opisującej pojedynczą klatkę obrazu systemu wizyjnego może być dokonana w sposób jeszcze bardziej zbliżony do jej biologicznego odpowiednika. Siatkówka ludzkiego oka posiada tzw. plamkę żółtą (miejsce największego upakowania czopków), gdzie jakość/rozdzielczość obrazu jest najlepsza. Właśnie takie podejście, w połączeniu z niedawno powstałą ideą sieci HTM (ang. Hierarchical Temporal Memory) może w przyszłości umożliwić stworzenie robotów prawdziwie „inteligentnych” – operujących abstrakcyjnymi pojęciami rozumianymi podobnie jak rozumie je człowiek.

Wspomniany pomysł reprezentacji obrazu pozyskiwanego z systemu wizyjnego został zaimplementowany przy użyciu dyskretnej transformaty falkowej o zmiennym (w przestrzeni dwuwymiarowej) progu (ang. threshold). Wartość wspomnianego współczynnika w danym punkcie ustalana jest w oparciu o jego odległość od tzw. „wirtualnej plamki żółtej” (punktu zainteresowania określonego w sprzężeniu zwrotnym przez algorytm) [2]. Tej odległości użyto jako jednego z parametrów funkcji Gaussa wyznaczającej wartość progu w punkcie. Dzięki temu możliwa stała się znacząca redukcja

objętości danych opisujących obraz, przy jednoczesnym zachowaniu najwyższej możliwej jakości w otoczeniu punktu zainteresowania (Rys. 1).



*Rys. 1. Krzywe pokazujące wartość progu wzdłuż osi x i y obrazu „Lena”. Jakość obrazu specjalnie została zredukowana, aby pokazać meritum idei – otoczenie punktu zaznaczonego krzyżykiem jest zawsze w maksymalnej jakości.
Praca własna.*

W realizacji praktycznej zastosowano uproszczoną wersję kompresji RLE (ang. Run-Length Encoding), gdzie kodowana była tylko ilość zer (inne wartości nie były kodowane). W wyniku kompresji obraz wynikowy mógł zajmować np. sześciokrotnie mniej danych (Rys. 1), nadal będąc w pełni funkcjonalnym źródłem informacji dla sieci HTM (obserwowany przez „wirtualną płamkę żółtą” detal ma maksymalną jakość).

To rozwiązanie (po zastosowaniu lepszego algorytmu kompresji, np. entropy-coding) zostanie przez autora zastosowane w celu umożliwienia wykorzystania klastra komputerowego jako dodatkowej mocy obliczeniowej systemu wizyjnego robota mobilnego.

LITERATURA

- [1] Nałęcz M., Duch W., Korbicz J., Rutkowski L., Tadeusiewicz R., Biocybernetyka i Inżynieria Biomedyczna 2000 – Tom 6 – Sieci Neuronowe, Akademicka Oficyna Wydawnicza Exit, Warszawa, 2000.
- [2] Podpora M., Computer vision in parallel computing, Przegląd Elektrotechniczny, ISSN 0033-2097, nr 11/2007, str. 68-70.