

# Raport - Projekt semestralny

Katarzyna Dziegiel, Magdalena Szpor, Agnieszka Szymczuk

5 lutego 2016

## 1 Wstęp

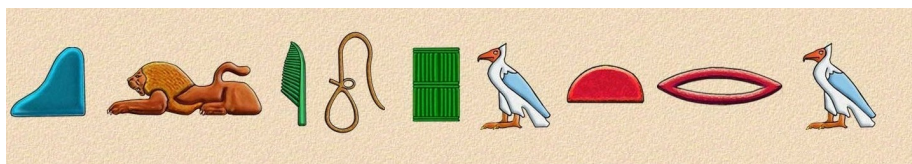
Celem projektu było napisanie programu rozpoznającego graficzne znaki – egipskie hieroglify, oraz tłumaczącego znalezione słowa na język polski.

## 2 Znaki egipskie

Pismo starożytnych Egipcjan opiera się na trzech rodzajach hieroglifów: znakach fonetycznych, znakach ideograficznych oraz determinatywach. Do wykonania projektu założono, że poszczególne znaki to pojedyncze litery lub znaki dwuzgłoskowe. W rzeczywistości mogą one oznaczać także pojedyncze słowa, a w połączeniu z innymi znakami przybierać inne znaczenie tworząc odpowiednie formy gramatyczne. Sposób odczytywania hieroglifów zależy od kontekstu ich występowania oraz układu. W projekcie przedstawiono 3 słowa oznaczające imiona egipskich władców: Chufu – znany powszechnie z tłumaczenia na język grecki jako Cheops, Kleopatra oraz Ptolemeusz. Imiona władców zapisywane były zazwyczaj jednoznacznie – jako litery (ewentualnie ich przymioty zapisywano w sposób symboliczny). Dla wyróżnienia ich w tekście imiona władców zapisywane były w tzw. kartuszach czyli owalnych „magicznych pętlach”.

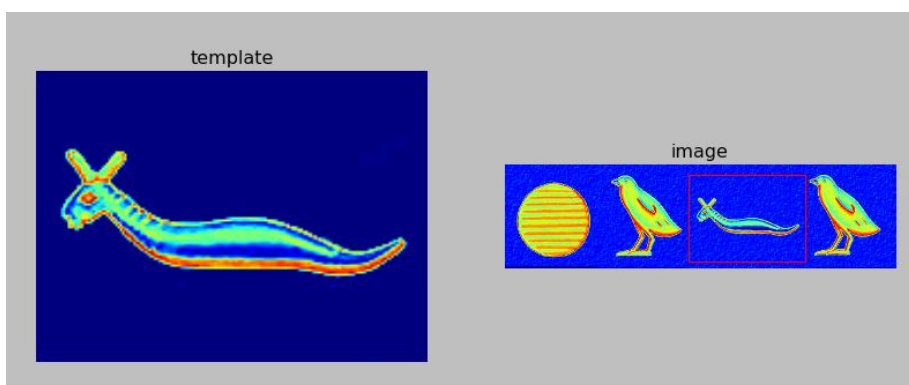
## 3 Zasada działania programu

Program napisany jest w języku Python i oparty jest głównie na funkcjach dostępnych w bibliotece scikit-image. Na początku zaczytywany jest rastrowy obraz przedstawiający ciąg znaków oznaczających jakieś słowo (rys.1). Ponieważ obraz przedstawia kolorowe, ciemne kształty na jasnym tle należało go najpierw zamienić na obraz czarno-biały, a następnie stworzyć negatyw.

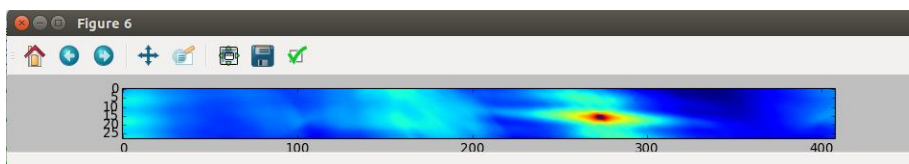


Rysunek 1: przykładowy napis

Litery wyszukiwane są na obrazie (rys. 2 image) za pomocą przygotowanych dla pojedynczych znaków wzorców (rys.2 template). Wzorec znajduje na obrazie obszary podobne do niego w stopniu ustalonym przez próg korelacji - czyli próg podobieństwa. (rys3)



Rysunek 2: Wzorec (z lewej), Wyszukowanie wzorca na obrazie (z prawej)



Rysunek 3: Mapa współczynnika korelacji

Ustalenie progu jest konieczne aby wybrane kształty nie były znajduwane też na fragmentach innych liter. Rys 3 przedstawia mapę współczynnika korelacji. Miejsca czerwone to obszary z największym podobieństwem, miejsca granatowe to obszary z zerowym podobieństwem, natomiast wszystkie pośrednie odcienie to podobieństwo zawierające się między skrajnymi wartościami. Po odnalezieniu litery wyświetlana jest jej nazwa. Gdy wszystkie wzorce zostaną sprawdzone litery sortowane są w kolejności w jakiej znajdują się na obrazie od lewej do prawej strony na zaczytanym obrazie i w ten sposób powstaje dane słowo. W celu

przetłumaczenia go na język polski używana jest funkcja słownika w którym zapisane zostało co jakie słowo oznacza. W ten sposób możliwe jest przetłumaczenie egipskich hieroglifów na słowa w języku polskim (rys.4).

```
Znalezione literki (PO SORTOWANIU) :  
['k', 'l', 'i', 'o', 'p', 'a', 't', 'r', 'a']  
  
Po egipsku:  
kliopatra  
  
Po polsku:  
Kleopatra  
>>>
```

Rysunek 4: Sortowanie liter i tłumaczenie wyrazów

## 4 Wnioski i perspektywa rozwoju programu

Program można wykorzystać do zamieniania wszelkich obrazów rastrowych na tekst pisany. Istnieje wiele możliwości rozbudowywania programu np. poprzez dodawanie kolejnych symboli lub nowych słów do słownika. W przyszłości można by dopracować tłumaczenie słów nie tylko na litery, ale na całe słowa i stworzyć bazę wyrazów które w określonej konfiguracji tworzą bardziej złożone struktury gramatyczne. Można też popracować nad wstępną analizą obrazu i za pomocą odpowiednich operacji morfologicznych modyfikować obraz aby precyzja wyszukiwania kształtów była jeszcze większa.