

**Dopasowanie obrazów cyfrowych – algorytm FFT****Streszczenie**

W pracy przedstawiono wyniki zastosowania bardzo efektywnego algorytmu implementującego dyskretne przekształcenie Fouriera (ang. Discrete Fourier Transform – DFT) znanego pod nazwą szybkie przekształcenie Fouriera (ang. Fast Fourier Transform - FFT). Omówiono jego wykorzystanie w podstawie analizy i przetwarzaniu sygnałów dwuwymiarowych – dopasowaniu do siebie obrazów cyfrowych (ang. image registration). W celu detekcji zmiany położenia geometrycznego jednego obrazu względem drugiego wykorzystano m.in. transformacje układów współrzędnych. Wkładem własnym autora w głównej mierze jest zaprezentowanie wizualizacji i omówienie wyników przeprowadzonej analizy problemu w części praktycznej projektu.

**Słowa kluczowe:** Dopasowanie obrazów cyfrowych, FFT, detekcja zmiany położenia.

**1. Wprowadzenie**

Dopasowanie obrazów cyfrowych jest zagadnieniem zawartym w dziedzinie przetwarzania obrazów. Jego opracowania naukowe zawarte są m.in. w wybranych pozycjach z literatury ([1],[5]-[7]). W zastosowaniach przetwarzania obrazów może zachodzić konieczność dopasowania do siebie dwóch obrazów przedstawiających tę samą scenę lub jej fragment, lecz po transformacji geometrycznej (jeden z obrazów względem drugiego może być min. przesunięty, obrocony, przeskalowany). Przykładowe zastosowanie operacji dopasowania obrazów znajdziemy w monitoringu wizyjnym, w systemach diagnostyki stanu maszyn czy budynków oraz rekonstrukcji map terenu wykorzystujących zdjęcia wykonane z „lotu ptaka” dzięki możliwości fotografowania z samolotu. Inną dziedziną nauki wykorzystującą omawiane zagadnienie może być m.in. medycyna (diagnostyka medyczna), w której porównywane są obrazy endoskopowe i z tomografii komputerowej. Niniejsza praca jest

prezentacją testów przy użyciu środowiska Matlab i omówieniem rezultatów.

**2. Szybkie przekształcenie Fouriera**

Szczegółowy opis teoretyczny na temat DFT oraz FFT dostępny jest w [2]. Dwuwymiarowa Transformata Fouriera (DFT2) obrazu I przedstawiona jest następująco (1):

$$F(i, k) = \sum_{m=0}^{M-1} \left( \sum_{n=0}^{N-1} I(m, n) e^{-\frac{2j\pi nk}{N}} \right) e^{-\frac{2j\pi mi}{M}} \quad (1)$$

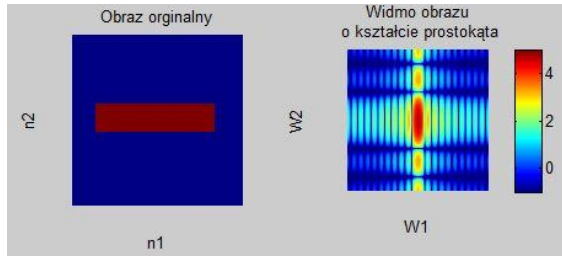
Znaczna część własności TF dla sygnałów w wymiarze jednowymiarowym jest również zachowana dla przypadku dwuwymiarowego, m.in. liniowość, przesunięcie w czasie, przesunięcie w częstotliwości, odwrócenie w czasie, różniczkowanie częstotliwości, spłot w czasie, spłot w częstotliwości oraz zachowanie energii. Cechą sygnałów dwuwymiarowych jest występowanie dodatkowej własności dla ciągów separowalnych. Jeśli sygnał dwuwymiarowy można przedstawić w postaci iloczynu zewnętrznego (ang. outer product) ciągów jednowymiarowych :

$$x[n_1, n_2] = x_1[n_1]x_2[n_2] = \begin{bmatrix} x_1[1] \\ x_1[2] \\ \dots \end{bmatrix} \begin{bmatrix} x_2[1] & x_2[2] & \dots \end{bmatrix} = \begin{bmatrix} x_1[1]x_2[1] & x_1[1]x_2[2] & \dots \\ x_1[2]x_2[1] & x_1[2]x_2[2] & \dots \\ \dots & \dots & \dots \end{bmatrix} \quad (2)$$

to jest to tzw. sygnał separowalny. Widmo sygnału separowalnego jest iloczynem zewnętrznych widm sygnałów jednowymiarowych.

$$X(\omega_1, \omega_2) = X_1(\omega_1)X_2(\omega_2) \quad (3)$$

Przykład sygnału separowanego przedstawiono na rys.1. Widmo obrazu o kształcie prostokąta, widmo to ma również kształt prostokątny.



Rys.1. Obraz o kształcie prostokątnym oraz widmo tego obrazu

Opis matematyczny obrazu  $F_2$  względem  $F_1$  (4):

$$F_2(w1, w2) = \sum_{m=0}^{M-1} \left( \sum_{n=0}^{N-1} F_1(n1, n2) e^{-\frac{2j\pi n w2}{N}} \right) e^{-\frac{2j\pi m w1}{N}}$$

Rysunek transformaty z rys.1 pokazuje większą energię dla dużych wertykalnych częstotliwości niż dla dużych horyzontalnych częstotliwości. Wynika to z faktu, iż na oryginalnym rysunku szybsze zmiany następują w przekroju pionowym (tzn., jest większa częstotliwość zmian) dlatego że prostokąt jest węższy w pionie.

Poniżej omówiono zagadnienia związane z analizą algorytmu dopasowania obrazów cyfrowych. W metodzie dopasowania obrazów skorzystano z właściwości przesunięcia przekształcenia całkowego, które jest zdefiniowane jako:

$$f_2(x, y) = f_1(x - x_0, y - y_0) \rightarrow$$

$$F_2(f_x, f_y) = e^{-j\pi(f_x x_0 + f_y y_0)} F_1(f_x, f_y) \quad (5)$$

gdzie:

$f_1(x, y), f_2(x, y)$  oznaczają dyskretne obrazy 2D.  $F_1(f_x, f_y), F_2(f_x, f_y)$  – oznaczają dwuwymiarowe transformaty Fouriera tych obrazów.

Przesunięcie  $(x_0, y_0)$  obrazu 2 względem obrazu 1 wyznaczane jest zgodnie z zależnością:

$$Fourier^{-1} \left( \frac{F_2(f_x, f_y) F_1^*(f_x, f_y)}{F_2(f_x, f_y) F_1^*(f_x, f_y)} \right) = Fourier^{-1} \left( e^{-j\pi(f_x x_0 + f_y y_0)} \right) = \delta(x - x_0, y - y_0) \quad (6)$$

Wynikiem zastosowania odwrotnej TF unormowanego widma wzajemnego obu obrazów jest impuls Diraca. Impuls ten jest przesunięty do punktu  $(x_0, y_0)$  w ciągłej przestrzeni kartezjańskiej. W celu detekcji przesunięcia wystarczy odczytać współrzędne maksimum impulsu.

Przy równoczesnym obrocie  $(\Theta)$  i przesunięciu  $(x_0, y_0)$  obrazów definicja (5) wyprowadzana jest do następującej postaci:

$$f_2(x, y) = f_1(x \cos \Theta + y \sin \Theta - x_0, -x \sin \Theta + y \cos \Theta - y_0) \quad (7)$$

$$F_2(f_x, f_y) = A \cdot F_1(f_x \cos \Theta + f_y \sin \Theta, -f_x \sin \Theta + f_y \cos \Theta) \quad (8)$$

Gdzie:  $A = e^{-j\pi(f_x x_0 + f_y y_0)}$ .

Po wykonaniu operacji zamiany współrzędnych kartezjańskich na biegunowe oraz po obliczeniu modułów obu stron równania rotacja (7) zostaje zdefiniowana za pomocą przesunięcia wzdłuż osi kątowej. Detekcja tego przesunięcia na podstawie (6) w układzie log-polar :

$$|F_2(r, \varphi)| = |F_1(r, \varphi - \Theta)| \quad (9)$$

Zmiana widma wynikająca z przeskalowania obrazu  $f_1(x, y)$  współczynnikiem  $a$  jest następującej postaci:

$$f_2(x, y) = f_1(ax, ay) \quad (9)$$

$$F_2(f_x, f_y) = \frac{1}{a^2} F_1\left(\frac{f_x}{a}, \frac{f_y}{a}\right) \quad (10)$$

Dzięki zlogarytmowaniu osi sprowadzono skalowanie do przesunięcia, które zgodnie z (6) wynosi:

$$F_2(\log f_x, \log f_y) = \frac{1}{a^2} F_1(\log f_x + \log a, \log f_y - \log a) \quad (11)$$

### 3. Algorytm dopasowania obrazów

Kolejność działań wykonawczych, wykorzystujących procedurę FFT w celu dopasowania dwóch obrazów, można podzielić na pięć wyróżniających się etapów: **1.** Obliczenie modułów FFT obu obrazów. W celu zredukowania niepożądanego przez nas zjawiska przecieku widmowego, które dotyczy DFT jak również FFT, należy zastosować okienkowanie (ang. windowing). Przeciek widma powoduje, że wyniki DFT stanowią jedynie aproksymacje rzeczywistych widm oryginalnych sygnałów wejściowych poddawanych próbkowaniu. Przeciek skutkuje tym, że dowolny sygnał wejściowy, którego częstotliwość nie jest dokładnie równa częstotliwości dla której wyznaczany jest dany prążek DFT, energia przecieka do wszystkich innych wyznaczanych prążków DFT. Wobec tego w celu minimalizacji przecieku widma na obrazy nakładane są dopasowane okna dwuwymiarowe;

m.in. spośród okien Hanninga, Hamminga lub Gaussa, następnie zostają obliczane dwuwymiarowe moduły FFT obrazów.

**2.** Zastosowanie transformacji modułów widm z układu kartezjańskiego do układu log-polar i obliczenie ich reprezentacji. Oczekiwanym efektem przeprowadzonej transformacji jest jednoczesna detekcja rotacji i skalowania obrazu. Dla obrazów dyskretnych współrzędne w układzie kartezjańskim reprezentują położenie piksela. Płaszczyzn kartezjańska i log-polar nie są próbkowane w tych samych pozycjach, wobec czego określany jest sposób próbkowania płaszczyzny log-polar (przyrosty kąta i logarytmu promienia) i szukany jest odpowiednik w układzie kartezjańskim dla każdego piksela z układu log-polar. Problem zmiennoprzecinkowych indeksów dla współrzędnych reprezentujących położenie pikseli w układzie kartezjańskim zostaje rozwiązany po użyciu m.in. interpolacji dwuliniowej. Interpolacja dwuliniowa jest opisana następującym równaniem:

$$M(x, y) = M_{k,l}(1-t)(1-u) + M_{k+1,l}(1-u)t + M_{k,l+1}(1-t)u + M_{k+1,l+1}tu \quad (12)$$

gdzie  $M(x,y)$  reprezentuje obliczoną jasność punktu obrazu o współrzędnych  $(x,y)$ , przy czym  $(x,y)$  oznaczają współrzędne przeliczone z układu log-polar do układu  $x-y$ .  $M_{p,q}$  reprezentują intensywność czterech pikseli, najbliższych sąsiadów punktu  $(x,y)$ , 't' jest częścią ułamkową 'x', natomiast 'u' jest częścią ułamkową 'y'.

**3.** Obliczenie wartości przesunięcia obrazu nr.2 w stosunku do obrazu nr.1 wzdłuż osi kątów i  $\log(r)$ .

**4.** Kolejnym etapem jest wykonanie przekształcania odwrotnego (6) do jednego z obrazów.

**5.** Ostatnim etapem procedury postępowania jest wyznaczenie przesunięcia zgodnie z zdefiniowanym odwrotnym przekształceniem (6).

## 4. Wizualizacja

W pracy projektowej wykorzystano obraz monochromatyczny o wymiarach szerokości i wysokości równych 512 pikseli. Rotacje wykonano funkcją 'imrotate' obracającą obraz względem punktu centralnego o zadany kąt przeciwnie do ruchu wskazówek zegara. Wyliczenie stałoprzecinkowych indeksów  $x-y$  z wartości zmiennoprzecinkowych, powstałych w wyniku

zastosowania zamiany współrzędnych kartezjańskich na log-polar, uzyskano za pomocą interpolacji dwuliniowej ('bilinear'). Skalowanie obrazu ('imresize') powoduje zmianę macierzowej wielkości obrazu (podobnie jak rotacja). Nowopowstałe pomiędzy poprzednimi pikselami szacowane elementy macierzy obrazu są obliczane z wykorzystaniem interpolacji dwuliniowej. Po przeprowadzeniu tych operacji należy wrócić do początkowej wielkości macierzowej obrazu. Wykonując operacje przesunięcia nie zmieniamy wielkości macierzowej obrazu. W przesunięciu obrazu o zadaną wartość pikseli tworzymy nową macierz obrazu z dwóch macierzy wartości: pierwszą przypisując macierz zerową o zadane przesunięcie oraz drugą z poprzednimi wartościami obrazu. Efektem takiego działania jest utrata danych względem obrazu oryginalnego. Okno dwuwymiarowe dla obrazu tworzymy przemnażając obraz przez iloczyn okien jednowymiarowych kolumn i wierszy macierzy obrazu. Wyniki powyższego postępowania przedstawiono na rys.1. Opis matematyczny obrazu  $I_2$  względem  $I_1$ :

$$\begin{bmatrix} I_2(x) \\ I_2(y) \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} I_1(x) \\ I_1(y) \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_{xy} \sin \Theta + \sigma_y \cos \Theta & \sigma_{xy} \cos \Theta - \sigma_x \sin \Theta & t_x \\ \sigma_y \sin \Theta & \sigma_y \cos \Theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} I_1(x) \\ I_1(y) \\ 1 \end{bmatrix} \quad (13)$$

$\sigma_x$  – przeskalowanie wzdłuż osi X,  
 $\sigma_y$  – przeskalowanie wzdłuż osi y,  $\sigma_{xy}$  – skręcenie,  
 $\Theta$  – rotacja względem początku układu współrzędnych,  
 $t_x$  – przesunięcie wzdłuż osi X,  $t_y$  – przesunięcie wzdłuż osi Y.

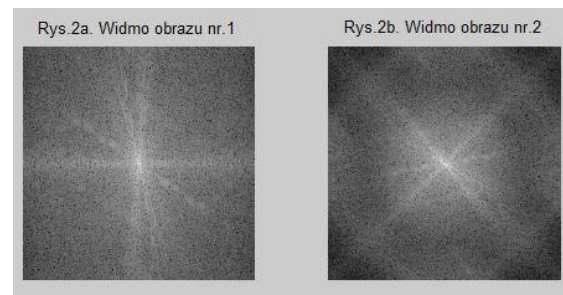
Przyjęto przekształcenie obrazu  $I_2$  zgodne z równaniem (13):

$$\begin{bmatrix} I_2(x) \\ I_2(y) \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \cdot \sin 45^\circ + 1.5 \cos 45^\circ & 0 \cdot \cos \Theta - 1.5 \sin 45^\circ & 12 \\ 1.5 \sin 45^\circ & 1.5 \cos 45^\circ & -8 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} I_1(x) \\ I_1(y) \\ 1 \end{bmatrix} \quad (14)$$



Rys.1 Obrazy z nałożonym oknem 2D Hanninga. Obraz nr.2 jest przeskalowany 1,5 razy, obrócony wokół środka o 45 stopni przesunięty w prawo oraz w dół odpowiednio 12 i 8 pikseli względem obrazu nr.1

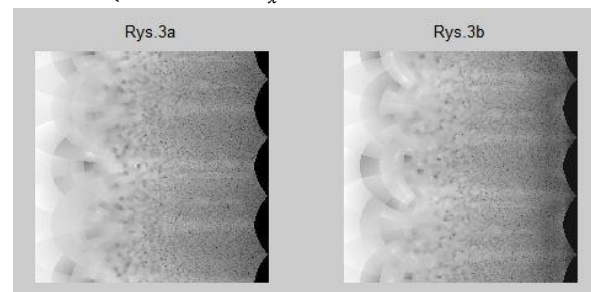
Zakres dużych zmian wartości modułów widm lepiej obrazują logarytmy wyznaczonych modułów. Wobec tego obliczono dwuwymiarową DFT z wykorzystaniem algorytmu FFT ('fft2'). Zlogarytmowane wartości DFT w celu wizualizacji transformacji Fouriera można uporządkować funkcją 'fftshift', która po przekątnej zamienia miejscami cztery podmacierze macierzy obrazu. Wynik operacji prezentuje rys.2.



Rys.2 Moduły widm dopasowywanych obrazów w skali logarytmicznej

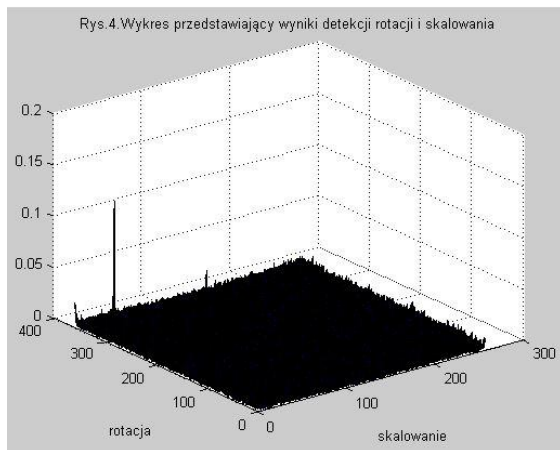
Moduły FFT obrazów nr.1 i nr.2 transformowane do układu log-polar, które w kolejnym kroku posłużą do pomiaru przesunięcia, zaprezentowano na rys.3. W implementacji zamiany układu współrzędnych kartezjańskich na log-polar należy wyznaczyć maksymalny zakres wektora wodzącego zgodnie z (15) oraz określić wartość kąta obrotu  $\phi$ . Następnie tworzona jest zerowa macierz o wymiarze wierszy równym długości wektora kąta obrotu i kolumn równym długości wektora promieni. Przypisanie wartości nowo powstałej macierzy następuje dzięki zastosowaniu algorytmu zależności matematycznej układów:

$$\begin{cases} r = e^{\rho}, \rho = \log \sqrt{x^2 + y^2} \\ \theta = \tan^{-1} \frac{y}{x} \text{ dla } x > 0 \end{cases} \quad (15)$$

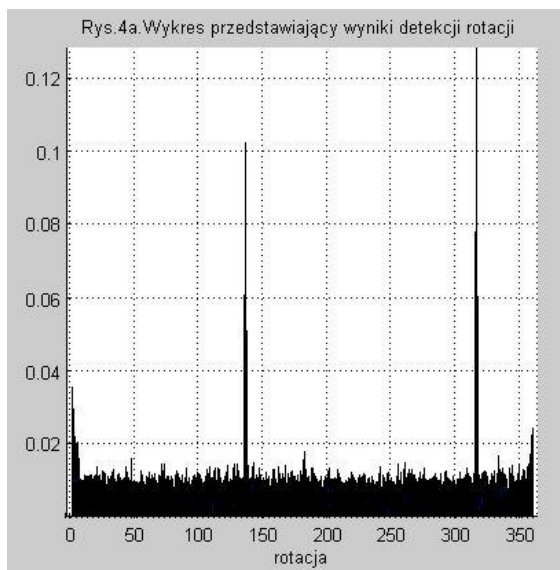


Rys.3 Reprezentacja modułów widm w układzie log-polar

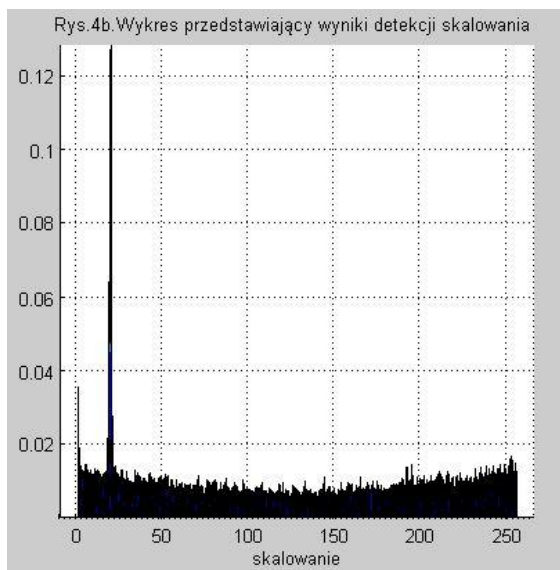
W celu pomiaru przesunięcia należy obliczyć 2d DFT, tym razem w układzie log-polar. Następnie zgodnie z (6) obliczono odwrotną TF z ilorazu iloczynu DFT obrazu nr.2 i sprzężonych wartości DFT obrazu nr.1 z iloczynem modułów DFT obu obrazów. Wizualizacje obliczeń uzyskano za pomocą funkcji 'surf', przedstawionych na rys.4, 4a, 4b.



Rys.4 Wyniki detekcji rotacji i skalowania



Rys.4a Wynik detekcji rotacji



Rys.4b Wynik detekcji skalowania

Wynikiem operacji IDFT są impulsy Diraca. Należy odczytać współrzędne jego maksimum (funkcja 'max'). Po przeskalowaniu osi  $\log r$  na  $r$ ,

wykorzystując przekształcenie (6), odczytano współczynnik skalowania. Wynikiem wyznaczania kąta rotacji  $\theta$  jest detekcja dwóch kątów, tj.  $\theta$  i  $(\theta+180)$ . Aby sprawdzić, która wartość kąta jest prawidłowa, wykonano odwrócenie tych rotacji i sprawdzenie, dla której wartość przesunięcia w osi kątowej wynosi zero. Następnie w wyniku przeprowadzenia przekształcenia odwrotnego (6) obrazu 2, wykorzystując transformację współrzędnych z układu log-polar do układu kartezjańskiego na podstawie równości:

$$\begin{cases} x = e^{\rho} \cos \theta \\ y = e^{\rho} \sin \theta \end{cases} \quad (16)$$

oraz korzystając z funkcji odwrotnej DFT 'ifft' otrzymano obraz ilustrowany na Rys.5. W algorytmie wykorzystano łącznie sześć transformat FFT oraz trzy odwrotne transformaty IFFT.



Rys.5. Etap końcowy dopasowania obrazu nr.2 względem obrazu nr.1

## 5. Podsumowanie

Zaprezentowana metoda dopasowania obrazów cyfrowych bazuje na przekształceniu Fouriera. Największą zaletą metody Fouriera - Mellina w porównaniu z klasycznymi metodami obszarowymi (min. bazujących na wyznaczeniu korelacji wzajemnej między obrazami) jest szybkość wykonania algorytmu. Zaletą metody jest stały czas wyznaczania parametrów dopasowania, zależny od wielkości obrazów. Opisana metoda wykorzystuje właściwość przekształcenia Fouriera w układzie współrzędnych log-polar polegającą na fakcie, że

skalowanie i rotacja obrazów powodują jedynie translacje widma ich obrazu w układzie log-polar. Wartość kąta obrotu oraz skali dopasowywanych obrazów wyznacza się analogicznie jak wartość przesunięcia.

## 6. Literatura

[1] Tomasz P. Zieliński: Cyfrowe przetwarzanie sygnałów, *Wydawnictwa Komunikacji i Łączności*, Warszawa 2005.

[2] Richard G. Lyons: Wprowadzenie do cyfrowego przetwarzania sygnałów, *Wydawnictwa Komunikacji i Łączności*, Warszawa 2010.

[3] Craig Marven, Gallian Ewers: Zarys cyfrowego przetwarzania sygnałów, *Wydawnictwa Komunikacji i Łączności*, Warszawa 1999.

[4] Krzysztof Duda: Analiza sygnałów biomedycznych, *Wydawnictwa AGH*, Kraków 2010.

[5] Jan Klimaszewski, Barbara Putz: Blokowy algorytm dopasowania obrazów multimodalnych w systemach monitorowania „Problemy eksploatacji. Maintenance Problems” , tom 81, numer 2/2011, 2011, s. 87-98.

[6] Ryszard Tadeusiewicz, Przemysław Korohoda: Komputerowa analiza i przetwarzanie obrazów, *Wydawnictwo Fundacji Postępu Telekomunikacji*, Kraków 1997.

[7] Zygmunt Wróbel, Robert Koprowski: Praktyka przetwarzania obrazów z zadaniami w programie Matlab, *Wydawnictwo EXIT*, Warszawa 2008.

[8]  
<<http://freesourcecode.net/matlabprojects/57372/foriermellinimageregistrationinmatlab#.WGLd2VPhDIU>>