

Maciej PHILIPP, Agnieszka TOMAKA
Marcin WRZEŚNIEWSKI

DETEKCJA I ESTYMACJA PARAMETRÓW RUCHU Z WYKORZYSTANIEM KLASTERYZACJI

Streszczenie. Opracowanie przedstawia metodę detekcji i estymacji parametrów ruchu, wykorzystującą algorytm klasteryzacji oraz podaje modyfikację procesu pasowania klastarów. Zaletą metody jest względnie mała wrażliwość na zniekształcenia sceny w ruchomych obrazach. Skuteczność metody została przedstawiona na podstawie analizy sekwencji obrazów o różnym stopniu złożoności. Wyniki przeprowadzonych badań zamieszczono w tabelach. Aby uwydatnić cel modyfikowania algorytmu, porównano wyniki detekcji ruchu na tych samych sekwencjach obrazów.

MOTION DETECTION AND ESTIMATION VIA CLUSTER MATCHING

Summary. The paper presents a routine of motion detection and estimation via cluster matching. Additionally the modification of cluster matching process is enclosed. Its strength consists in little sensitivity to scene deformation in moving images. The efficiency of this routine has been presented in a tabular form on the ground of image sequences' analysis for different complexity levels. To emphasise the purpose of modifications, the results of motion detection for the same image sequences have been compared.

1. Wprowadzenie

Jednym z zastosowań komputerowych systemów wizyjnych jest detekcja i estymacja parametrów ruchu obiektów sceny na podstawie zmian jasnościowych w obrazach należących do sekwencji. Stosuje się do tego celu między innymi metody obliczające przepływ optyczny (*optical flow*) [2],[3],[1],[7]. Polegają one na odnajdywaniu odpowiedniości pomiędzy obrazami z sekwencji w wybranych punktach kolejnych ramek (proces ten nazywany pasowaniem). Ruch reprezentowany jest jako prostopadły rzut wektorów prędkości punktów przestrzeni trójwymiarowej na płaszczyznę obrazu (płaszczyznę odwzorowania). Reprezentacja

przepływu optycznego może mieć postać wektorów rozmieszczonych na równomiernej siatce bądź wektorów ułożonych nierównomiernie, tzn. w punktach charakterystycznych (np. na krawędziach, w wierzchołkach) [3],[4]. Wybór najskuteczniejszej techniki pasowania zależy od struktury obrazu. Wiąże się to z wieloma problemami znacząco utrudniającymi wykrywanie obiektów. Można do nich zaliczyć: zmiany oświetlenia sceny w czasie, zmiany kształtów oraz sztywności ruchomych obiektów w wyniku wykonywania przez nie ruchu rotacyjnego lub prostopadłego do detektora, przesłonięcia w scenach o zróżnicowanym tle, ruchy elementów przezroczystych i półprzezroczystych.

Algorytm detekcji, który został przedstawiony w [1], różni się od typowych metod wykrywania ruchu. Pasowaniu podlegają spójne jasnościowo obszary obrazów (obiektów), które uzyskuje się w procesie klasteryzacji (stąd też obszary te nazwano klasterymi¹). Obiekty w ruchu są reprezentowane przez klasterzy przesunięte względem siebie, które posiadają zbliżone cechy jasnościowe i cechy kształtu.

Metoda ta jest interesująca co najmniej z dwóch powodów:

- ze względu na możliwość generacji regularniejszego przepływu optycznego niż w przypadku innych metod. W [1] jest to pokazane na podstawie eksperymentów, w których porównywano tę metodę z metodą gradientową i afiniczną;
- ze względu na możliwość wykrywania ruchu przy dużym zakresie zmian położenia obiektów w obrazie, zmian jasnościowych i deformacji ich kształtu.

Poniżej przedstawiono wyniki badań efektywności wykorzystania algorytmu z [1] do pasowania klastery przy dużym zakresie ruchu obiektu na scenie, przy występowaniu większej liczby obiektów oraz ich deformacji. Detekcja ruchu jest różna dla każdej sceny, gdyż zależy od warunków pozyskiwania sekwencji obrazów, charakteru ruchu, rodzaju i liczby ruchomych obiektów. Proces optymalizacji wag zastosowany w algorytmie powoduje, że przy małym ruchu elementów występuje większa tolerancja na zmiany ich kształtów i jasności w kolejnych obrazach. Gdy odległość pomiędzy elementami zwiększa się, tolerancja ta musi maleć, by pasowanie odpowiednich obiektów było właściwe.

W przypadku gdy zakres ruchu elementów jest duży i ruch przebiega w sposób nieuporządkowany, metoda zawodzi nawet wtedy, gdy kształty obiektów i ich jasności są nienaruszone. Problem ten występuje również wówczas, gdy kształty i jasności obiektów sceny są do siebie zbliżone. Stąd też proponuje się modyfikację metody, która wykrywa obiekty o znacznych zmianach położenia poprzez wykorzystanie w procesie optymalizacji wag większej liczby cech niosących informacje o kształcie. Poniżej przedstawiono proces pasowania sekwencji obrazów w celu uwydatnienia problemów związanych z jego zastosowaniem do detekcji i estymacji parametrów ruchu. zilustrowano również wyniki pasowania wykorzystując zaproponowaną modyfikację metody.

¹ Klaster - grupa punktów spójnych obszarowo i jakościowo.

W punkcie drugim przedstawiono opis kolejnych etapów algorytmu, w punkcie trzecim opis przeprowadzonych eksperymentów, natomiast w punkcie czwartym przedstawiono modyfikację metody oraz wynikające z niej korzyści. W zakończeniu podano wnioski i sugestie odnośnie do dalszych prac.

2. Metoda wykrywania ruchu poprzez pasowanie klastrów

Przedstawiony w [1] algorytm rozpoznawania ruchu zawiera trzy podstawowe etapy. Etap pierwszy polega na podziale obrazu na grupy punktów o podobnych cechach. Przestrzeń cech jest trójwymiarowa: pozycja (współrzędne x i y punktu w obrazie) i jasność. Grupy tych punktów nazywamy klastrami, a sam proces podziału - procesem klasteryzacji obrazu. Drugi etap polega na pasowaniu klastrów dwóch obrazów w wyniku zastosowania iteracyjnego procesu minimalizacji odległości euklidesowej, obliczanej na podstawie rozszerzonego wektora cech klastrów. Etap trzeci dotyczy znajdowania klastrów reprezentujących ruchome obiekty sceny.

OZNACZENIA:

Punkty obrazu oznaczono jako p^i , gdzie $i \in \{1, 2, \dots, N\}$; N - liczba punktów obrazu. Każdy punkt opisano za pomocą trzech wielkości (cech):

$$p^i = \begin{bmatrix} x & y & I(x,y) \end{bmatrix}^T,$$

gdzie:

x, y - współrzędne położenia punktu w kartezjańskim układzie współrzędnych,

$I(x,y)$ - funkcja jasności punktu,

$$\forall_{x,y \in N \cup \{0\}} I(x,y) \in \{2^n, n \in N \cup \{0\} \wedge n \leq 8\}.$$

Klaster oznaczono jako k^j , gdzie $j \in \{1, 2, \dots, K\}$; K - liczba klastrów.

Klaster opisano podobnie jak punkt obrazu jako wektor trzech cech:

$$m^j = \begin{bmatrix} m'_x & m'_y & m'_z \end{bmatrix},$$

gdzie:

$$m'_x = \frac{\sum x}{N_j}, \quad m'_y = \frac{\sum y}{N_j}, \quad m'_z = \frac{\sum I(x,y)}{N_j}, \quad (1.0)$$

gdzie:

m'_x - współrzędna środka ciężkości klastru k^j w osi X ,

m'_y - współrzędna środka ciężkości klastru k^j w osi Y ,

m'_z - średnia jasność punktów w klastrze k^j ,

N_j - liczba punktów w klastrze k^j .

Macierz W^j zdefiniowano jako macierz wag.

$$W^j = \begin{bmatrix} w'_x & 0 & 0 \\ 0 & w'_y & 0 \\ 0 & 0 & w'_z \end{bmatrix},$$

gdzie:

$$w'_x = \frac{c^j}{\sigma_x^j}, \quad w'_y = \frac{c^j}{\sigma_y^j}, \quad w'_z = \frac{c^j}{\sigma_z^j}, \quad c^j = (\sigma_x^j \sigma_y^j \sigma_z^j)^{\frac{1}{3}}, \quad (1.1)$$

$$(\sigma_x^j)^2 = \frac{1}{(N_j - 1)} \sum_{z \in k^j}^{N_j} (x - m'_x)^2, \quad (1.2-a)$$

$$(\sigma_y^j)^2 = \frac{1}{(N_j - 1)} \sum_{y \in k^j}^{N_j} (y - m'_y)^2, \quad (1.2-b)$$

$$(\sigma_z^j)^2 = \frac{1}{(N_j - 1)} \sum_{z \in k^j}^{N_j} (z - m'_z)^2, \quad (1.2-c)$$

$$z = I(x, y), \quad (1.2-d)$$

gdzie:

$(\sigma_x^j)^2$ - nieobciążona wariancja liczona dla współrzędnych x punktów z klastra k^j ,

$(\sigma_y^j)^2$ - nieobciążona wariancja liczona dla współrzędnych y punktów z klastra k^j ,

$(\sigma_z^j)^2$ - nieobciążona wariancja liczona dla współrzędnych z punktów z klastra k^j .

Punkt p^j jest przypisany do klastra k^j , dla takiego $j \in \{1, 2, \dots, K\}$, dla którego wyrażenie

$$(p^j - m^j)^T W^j (p^j - m^j) \quad (1.3)$$

osiąga wartość minimalną.

Macierz U zdefiniowano jako diagonalną macierz wag:

$$U = \begin{bmatrix} u_x & 0 & 0 & 0 & 0 \\ 0 & u_y & 0 & 0 & 0 \\ 0 & 0 & u_z & 0 & 0 \\ 0 & 0 & 0 & u_m & 0 \\ 0 & 0 & 0 & 0 & u_s \end{bmatrix}$$

Przez k^j , gdzie $j \in \{1, 2, \dots, K_t\}$, oznaczamy klaster obrazu przetworzonego w czasie t . Przez k^l , gdzie $l \in \{1, 2, \dots, K_{t+\Delta t}\}$, oznaczamy klaster obrazu przetworzonego w czasie $t+\Delta t$.

Klaster k^j spawany jest z klasterem k^l dla takiego $l \in \{1, 2, \dots, K_{t+\Delta t}\}$, dla którego wyrażenie

$$(b^j - c^l)^T U (b^j - c^l) \quad (1.4)$$

osiąga wartość minimalną.

Korzystając z metody mnożników Lagrange'a, minimalizuje się wyrażenie (1.4) przy ograniczeniu $u_x u_y u_z u_m u_s = 1$ otrzymując wzory pozwalające na obliczenie wartości wag:

$$u_y = \frac{u_x \sum_{j=1}^{k'} (m'_x - m'_j)^2}{\sum_{j=1}^{k'} (m'_y - m'_j)^2}, \quad (1.5-a)$$

$$u_z = \frac{u_x \sum_{j=1}^{k'} (m'_z - m'_j)^2}{\sum_{j=1}^{k'} (m'_z - m'_j)^2}, \quad u_m = \frac{u_x \sum_{j=1}^{k'} (m'_x - m'_j)^2}{\sum_{j=1}^{k'} (m'_m - m'_j)^2}, \quad u_s = \frac{u_x \sum_{j=1}^{k'} (m'_x - m'_j)^2}{\sum_{j=1}^{k'} (m'_s - m'_j)^2}, \quad (1.5-b)$$

gdzie:

$$u_x = \left[\sum_{j=1}^{k'} (m'_y - m'_j)^2 \right]^{\frac{1}{5}} \left[\sum_{j=1}^{k'} (m'_z - m'_j)^2 \right]^{\frac{1}{5}} \left[\sum_{j=1}^{k'} (m'_m - m'_j)^2 \right]^{\frac{1}{5}} \left[\sum_{j=1}^{k'} (m'_s - m'_j)^2 \right]^{\frac{1}{5}} \left[\sum_{j=1}^{k'} (m'_x - m'_j)^2 \right]^{\frac{4}{5}}, \quad (1.5-c)$$

Rozszerzone wektory cech dla klastrów k^j i k^l zdefiniowano jako:

$$b^j = \begin{bmatrix} m^j \\ m'_m \\ m'_c \end{bmatrix}, \quad c^l = \begin{bmatrix} m^l \\ m'_m \\ -m'_s \end{bmatrix},$$

gdzie:

$$m'_x = \frac{1}{N_j} \sum_{x,y \in k^j} \{I(x,y,t) - I(x,y,t + \Delta t)\}, \quad (1.6)$$

$$m'_m = \sum_{x,y \in k^j} [(x - \bar{x}^j)^2 - (y - \bar{y}^j)^2] I(x,y), \quad (1.7)$$

$$\bar{x}^j = \frac{\sum_{x,y \in k^j} x I(x,y)}{\sum_{x,y \in k^j} I(x,y)}, \quad \bar{y}^j = \frac{\sum_{x,y \in k^j} y I(x,y)}{\sum_{x,y \in k^j} I(x,y)}, \quad (1.8)$$

zaś:

m'_x - średnia z różnic jasności punktów w klastrze k^j ,

m'_m - suma kwadratowych momentów centralnych liczona dla współrzędnych x i y punktów z klastra k^j ,

m^j, m^l - wektory cech (m'_x, m'_y, m'_z) opisujące klastry k^j i k^l .

Podobnie liczy się wartości współrzędnych wektora c^l . Ujemny znak przed wartością średniej z różnic jasności punktów w klastrze k^l powoduje, że wartość tę liczy się ze wzoru:

$$m'_x = \frac{1}{N_l} \sum_{x,y \in k^l} \{I(x,y,t + \Delta t) - I(x,y,t)\}. \quad (1.9)$$

2.1. Algorytm klasteryzacji obrazu

DANE WEJŚCIOWE: zadana wartość określająca liczbę klastery K , zadana wartość określająca próg dokładności Tr , dwa kolejne obrazy z sekwencji.

DANE WYJŚCIOWE: mapa przedstawiająca rozkład klastery dla zadanych obrazów, wielkości opisujące klaster.

Krok 1. Podziel obraz na K kwadratowych regionów o jednakowym polu powierzchni. Przejdź do kroku 2.

Krok 2. Dla każdego klastera k^j , gdzie $j \in \{1, 2, \dots, K\}$, oblicz wartości $m'_x, m'_y, m'_z, (\sigma'_x)^2, (\sigma'_y)^2, (\sigma'_z)^2$ korzystając ze wzorów (1.0), (1.2-a) do (1.2-d). Przejdź do kroku 3.

Krok 3. Jeżeli $\sigma'_z = 0$, gdzie $j \in \{1, 2, \dots, K\}$, wartości wag z macierzy W^j ustaw jak poniżej i przejdź do kroku 5, jeżeli nie, przejdź do kroku 4.

$$w'_x = w'_y = 0,$$

$$w'_z = 1.$$

Krok 4. Dla każdego klastera k^j , gdzie $j \in \{1, 2, \dots, K\}$, oblicz wartości macierzy W^j . Przejdź do kroku 5.

Krok 5. Wyznacz, na podstawie kryterium ważonej kwadratowej odległości euklidesowej, przynależność każdego punktu obrazu do odpowiadającego mu klastera. Oblicz wartość wyrażenia (1.3) dla wszystkich par $(i, j) \in \{1, 2, \dots, N\} \times \{1, 2, \dots, K\}$ i każdemu (i) przypisz takie (j) , że

$$p'_o = \arg \min_{j \in \{1, 2, \dots, K\}} (p^i - m^j)^T W^j (p^i - m^j)$$

Przejdź do kroku 6.

Krok 6. Odszukaj klaster k^j , gdzie $j \in \{1, 2, \dots, K\}$, o maksymalnym przesunięciu środka ciężkości. Jeżeli wartość przesunięcia jest większa od zadanego progu Tr , przejdź do kroku 2, jeżeli nie, przejdź do kroku 7.

Krok 7. Jeżeli przetworzono dwa obrazy, to przejdź do kroku 8, jeżeli nie, to powtórz algorytm klasteryzacji dla następnego obrazu z sekwencji. Kolejne obrazy traktowane są niezależnie od siebie.

Krok 8. Koniec.

2.2. Algorytm pasowania klastery

DANE WEJŚCIOWE: dwa przetworzone w procesie klasteryzacji obrazy (przez k^j , gdzie $j \in \{1, 2, \dots, K\}$), oznaczamy klaster obrazu przetworzonego w czasie t ; przez k^l , gdzie $l \in \{1, 2, \dots, K_{t+\Delta t}\}$, oznaczamy klaster obrazu przetworzonego w czasie $t+\Delta t$), wielkości opisujące klaster.

DANE WYJŚCIOWE: pary odpowiadających sobie klastrów.

Krok 1. Dla każdego klastrera k^j , gdzie $j \in \{1, 2, \dots, K\}$, oblicz wartości m'_x, m'_m korzystając ze wzorów (1.6) do (1.9). Przejdź do kroku 2.

Krok 2. Ustaw początkowe wagi dla u_x, u_y, u_z, u_m, u_s .

Wagi sugerowane w [1] $u_x = u_y = u_z = u_s = 10, u_m = 10^{-4}$. Przejdź do kroku 3.

Krok 3. Wyznacz, na podstawie kryterium ważonej kwadratowej odległości euklidesowej, odpowiadające sobie klasterzy. Oblicz wartość wyrażenia (1.4) dla wszystkich par $(j, l) \in \{1, 2, \dots, K\} \times \{1, 2, \dots, K_{l+\Delta l}\}$ i każdemu (j) przypisz takie (l) , że

$$k'_a = \arg \min_{l \in \{1, 2, \dots, K_l\}} (b^j - c^l)^T U (b^j - c^l)$$

Przejdź do kroku 4.

Krok 4. Jeżeli wartość

$$D_T^n = D_T^{n+1},$$

gdzie:

$$D_T = \sum_{j=1}^K k'_a,$$

it - numer iteracji,

przejdź do kroku 6, jeżeli nie, przejdź do kroku 5.

Krok 5. Oblicz wartości wag u_x, u_y, u_z, u_m, u_s według wzorów (1.5-a) do (1.5-c) i przejdź do kroku 3.

Krok 6. Koniec.

2.3. Algorytm poszukiwania przemieszczeń

DANE WEJŚCIOWE: wielkości opisujące klasterzy, zadana wartość progu ufności α .

DANE WYJŚCIOWE: indeksy niestacjonarnych klastrów.

Krok 1. Dla każdego klastrera k^j , gdzie $j \in \{1, 2, \dots, K\}$, oblicz wartość

$$\gamma = \alpha \max_j |m'_s|.$$

Przejdź do kroku 2.

Krok 2. Jeżeli $\gamma < |m'_s|$, gdzie $j \in \{1, 2, \dots, K\}$, zapamiętaj indeks j (klaster niestacjonarny). Przejdź do kroku 3.

Krok 3. Koniec.

Uwaga:

Rząd średniej złożoności obliczeniowej całego algorytmu wynosi:

$$O(T) = O(n^2),$$

gdzie: n - liczba klastrów.

Złożoność pamięciowa zależy od liczby klastrów. Liczba bajtów pamięci potrzebna do przechowywania danych opisujących klasterzy jest wprost proporcjonalna do ich liczby:

$$\text{liczba_bajtów} = 204 * \text{liczba_klastrów}.$$

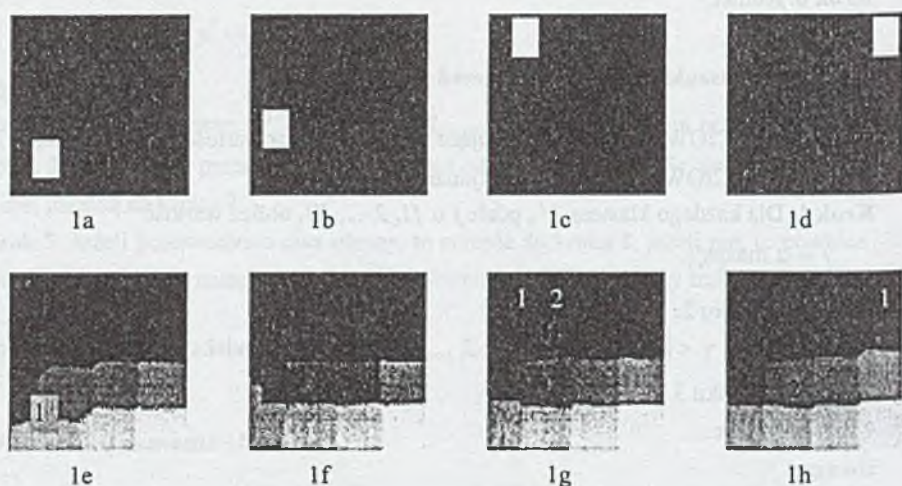
3. Eksperymenty

Celem badań była próba oceny efektywności działania metody przedstawionej w [1]. Do doświadczeń użyto komputerowej implementacji algorytmu. Aplikacja została napisana w języku C++. Jako narzędzie wybrano kompilator firmy BORLAND w wersji 3.1. Analiza dotyczyła wpływu zjawisk zakłócających na poprawność działania metody klasteryzacji obrazu. Te zjawiska zakłócające to:

- 1) duże przemieszczenie obiektu,
- 2) duża liczba ruchomych obiektów,
- 3) deformacja konturu obiektu.

Badania były wykonywane na obrazach testowych o rozdzielczości 256x256 pikseli w 256 stopniowej skali poziomów szarości.

Badaniom poddano sekwencje obrazów zawierające sztywne figury o jednorodnej jasności. Pierwsza sekwencja (rys.1) przedstawia ruch białego prostokąta na ciemnym tle. Przemieszczenie obiektu pomiędzy jego położeniem początkowym a kolejnym ulega zwiększeniu. Obrazy 1a, 1b, 1c, 1d zostały podzielone na 16 klastrow (1e, 1f, 1g, 1h). Obraz 1a pasowano kolejno z obrazem 1b, 1c i 1d. W każdym z przypadków nastąpiło prawidłowe wykrycie prostokąta. Na podstawie danych z tabeli 1 wnioskuje się, iż jest to wynik bardzo małych zmian parametrów m_z , m_m i m_s .



Rys. 1. Pierwsza sekwencja obrazów a, b, c, d i rozkład klastrow w nich zawartych e, f, g, h

Fig. 1. First sequence of images a, b, c, d and distribution of clusters e, f, g, h

Tabela 1

Spasowane klastery	Δm_x	Δm_y	$\Delta m_z / m_{zu}$	$\Delta m_m / m_{mu}$	$\Delta m_s / m_{su}$	D
1a[1]-1b[1]	13,93	42,74	0	-0,01	0,04	365896,70
1a[1]-1b[2]	-39,13	45,05	0,67	-0,41	1,06	173036705,15
1a[1]-1c[1]	-2	171	0	0	-0,01	223096,24
1a[1]-1c[2]	-44,55	175,45	0,66	-0,09	0,94	267700500,76
1a[1]-1d[1]	-176,19	169	0	0	-0,04	2755738,58
1a[1]-1d[2]	-42,73	42,55	0,7	-0,47	1,07	167483057,15

OZNACZENIA:

$$\Delta m_x = m'_x - m''_x,$$

$$\Delta m_y = m'_y - m''_y,$$

$$\Delta m_z / m_{zu} = (m'_z - m''_z) / m'_z,$$

$$\Delta m_m / m_{mu} = (m'_m - m''_m) / m'_m,$$

$$\Delta m_s / m_{su} = (m'_s - m''_s) / m'_s,$$

$$D = (b' - c')^T U (b' - c').$$

Pierwsza kolumna tabeli zawiera nazwy pasowanych ze sobą klasterów w obrazach. Dla klastra 1a[1], reprezentującego kwadrat na obrazie 1a, algorytm znalazł najlepszy odpowiednik na obrazie 1b w klastrze 1b[1] (klastery te, oznaczone numerem 1, pokazane są na obrazach 1e i 1f). Parametry tego spasowania znajdują się w pierwszym wierszu tabeli. Następnym najbliższym klastrzem dla 1a[1] na obrazie 1b jest klaster 1b[2] (klaster nr 2 na obrazie 1f), co widać w wierszu drugim. Wiersz trzeci i czwarty przedstawia tę samą sytuację dla obrazów 1a i 1c, a wiersz piąty i szósty dla obrazów 1a i 1d.

Stosunki odległości $\frac{D_{1a[1]-1b[1]}}{D_{1a[1]-1b[2]}} = 0,0021$, $\frac{D_{1a[1]-1c[1]}}{D_{1a[1]-1c[2]}} = 0,0008$, $\frac{D_{1a[1]-1d[1]}}{D_{1a[1]-1d[2]}} = 0,0016$, świadczą

o bardzo wysokim prawdopodobieństwie znalezienia odpowiadających sobie par klasterów. Sytuacja taka zachodzi w przypadku jednolitego tła i sztywnego obiektu o dużym kontraście. Każde dodatkowe zakłócenie w postaci niejednorodności tła lub obecności dodatkowych elementów zmniejsza prawdopodobieństwo prawidłowego wykrycia obiektu.

W tabeli 1.1 przedstawiono dane dotyczące procesu klasteryzacji obrazów 1a, 1b, 1c, 1d. W pierwszej kolumnie tabeli podano nazwy obrazów, dla których w kolumnach drugiej i trzeciej przedstawiono czasy trwania procesu klasteryzacji i liczby iteracji. Do podziału obrazów 1a, 1b, 1c, 1d użyto 16 klasterów.

Tabela 1.1

Nazwa obrazu	t [min.]	l
1a	8,4	9
1b	3,8	4
1c	7,6	8
1d	6,6	7

OZNACZENIA:

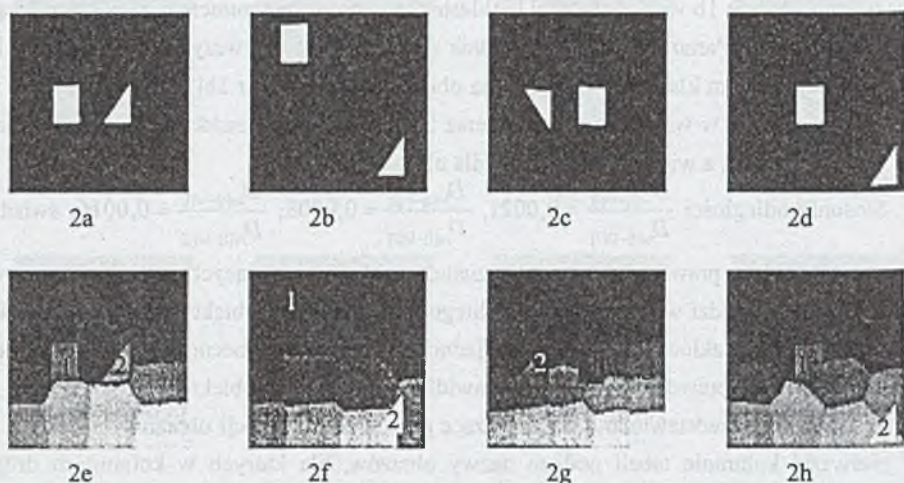
l - liczba iteracji,

t - czas trwania iteracji.

W następnej sekwencji obrazów (rys. 2) pojawił się dodatkowy element (trójkąt), który zwiększa złożoność sceny. Różni się on od kwadratu innym współczynnikiem kształtu.

Oba obiekty zmieniają położenie w stosunku do ich położenia na obrazie 2a.

Na obrazach 2b i 2c odległości pomiędzy kwadratami i trójkątami są zbliżone. Na obrazie 2c nastąpił dodatkowo efekt zamiany pozycji obiektów (trójkąt pojawił się na miejscu kwadratu i odwrotnie). Na obrazie 2d kwadrat pojawił się w połowie odległości między kwadratem i trójkątem w położeniu pierwotnym, a położenie trójkąta uległo znacznemu przesunięciu w kierunku prawego dolnego rogu obrazu.



Rys. 2. Druga sekwencja obrazów a, b, c, d i rozkład klasterów w nich zawartych e, f, g, h

Fig. 2. Second sequence of images a, b, c, d and distribution of clusters e, f, g, h

Tabela 2

Spasowane klastery	Δm_x	Δm_y	$\Delta m_z/m_{za}$	$\Delta m_m/m_{ma}$	$\Delta m_s/m_{sa}$	D
2a[1]-2b[1]	16,94	86,24	0	-0,01	0,03	401435,17
2a[1]-2b[2]	-128,93	-83,15	0,01	0,64	-0,04	7935570,42
2a[2]-2b[2]	-50,95	-74,81	0	0,13	-0,04	1365054,50
2a[2]-2b[1]	94,92	94,58	-0,01	-1,41	0,02	4679678,84
2a[1]-2c[1]	-68,92	-3,58	0	0,02	-0,12	130372,46
2a[1]-2c[2]	2,58	0,4	0	0,61	0,84	2789502,22
2a[2]-2c[2]	80,66	8,56	0	0,08	0,52	230099,59
2a[2]-2c[1]	9,16	4,58	-0,02	-1,33	-2,44	2477764,35
2a[1]-2d[1]	-37,1	-1,97	0	-0,01	0,04	39831,40
2a[1]-2d[2]	-148,66	-97,4	0,02	0,63	-0,09	18558354,15
2a[2]-2d[1]	40,88	6,38	-0,02	-1,41	0	521962,62
2a[2]-2d[2]	-70,68	-89,06	0	0,12	-0,14	15086748,07

W przypadku pasowania obrazów 2a z 2b i 2a z 2c nastąpiło poprawne wykrycie odpowiadających sobie figur. W przypadku pasowania obrazu 2a z 2d obie figury zostały spasowane z kwadratem, którego położenie na obrazie 2d, w przeciwieństwie do położenia trójkąta na tym obrazie, jest bliskie położeniom pierwotnym tych figur na obrazie 2a. Błędne spasowanie figur wynika więc z nierównomiernego oddalenia tych samych figur względem siebie. Szanse na spasowanie tych figur będą większe, jeśli poszczególne figury z obrazu różnić się będą między sobą jasnością. Problem błędnego spasowania udało się rozwiązać poprzez pasowanie odwrotne, tzn. obrazu 2d z 2a.

Tabela 2 przedstawia parametry pasowania klasterów. Pierwsza kolumna tabelki zawiera nazwy pasowanych ze sobą klasterów z kolejnych obrazów. Klaster 2a[1], reprezentujący kwadrat z obrazu 2a, został spasowany z klasterem 2b[1] z obrazu 2b (klastery te, oznaczone numerem 1, pokazane są na obrazach 2e i 2f). Wyniki tego spasowania znajdują się w pierwszym wierszu tabeli. Następnym najbliższym klasterem dla 2a[1] na obrazie 2b jest klaster 1b[2] (klaster nr 2 na obrazie 2f), co widać w wierszu drugim. Wiersz trzeci i czwarty przedstawiają tę samą sytuację dla klastera 2a[2] reprezentującego trójkąt. W tym przypadku okazało się, że kolejnym, najbliższym klasterem dla klastera przedstawiającego trójkąt jest klaster przedstawiający kwadrat. Kolejne cztery wiersze przedstawiają wyniki pasowania dla obrazów 1a i 1c. W przypadku pasowania klasterów obrazu 1a z 1d okazało się, że klaster 2a[2], reprezentujący trójkąt, został błędnie spasowany z klasterem 2d [1] reprezentującym kwadrat.

Stosunki odległości pasowanego kwadratu z obrazu 2a z odpowiednikami na obrazach z sekwencji wg tabeli 2 wynoszą:

$$\frac{D_{2a[1]-2a[1]}}{D_{2a[1]-2a[2]}} = 0,05, \quad \frac{D_{2a[1]-2a[1]}}{D_{2a[1]-2a[2]}} = 0,046, \quad \frac{D_{2a[1]-2a[1]}}{D_{2a[1]-2a[2]}} = 0,002$$

Podobnie dla trójkąta:

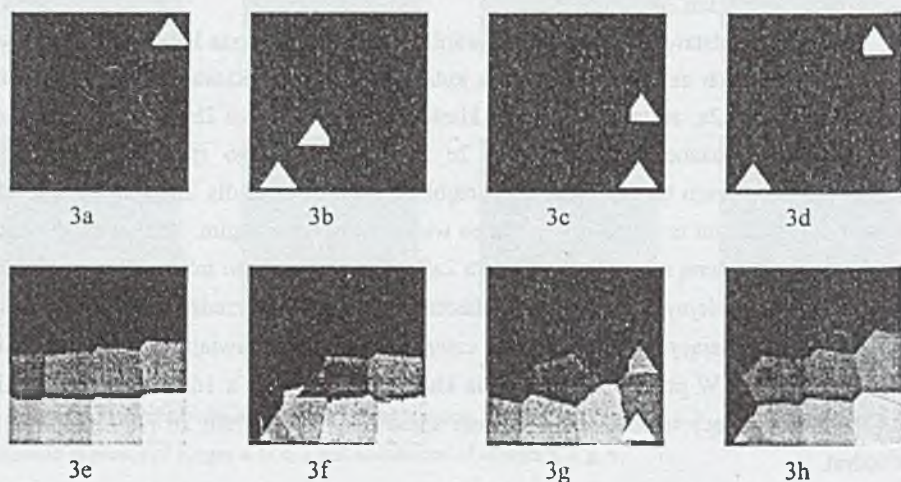
$$\frac{D_{2a[2]-2a[2]}}{D_{2a[2]-2a[1]}} = 0,29, \quad \frac{D_{2a[2]-2a[2]}}{D_{2a[2]-2a[1]}} = 0,09, \quad \frac{D_{2a[2]-2a[1]}}{D_{2a[2]-2a[2]}} = 0,03$$

W porównaniu z pierwszym eksperymentem stosunki odległości zwiększyły się co najmniej o rząd. Wynika to z pojawienia się dużych wartości $\Delta m_m / m_{ma}$ i $\Delta m_s / m_{sa}$. Analizując dane z tabeli 2 można wyciągnąć wniosek, iż dla obrazów 2b i 2c o poprawnym spasowaniu figur zadecydowało podobieństwo cech m_s , m_m , m_s . Gdy ruch figur jest nieuporządkowany, jak np. na obrazie 2d, na niekorzyść poprawnego spasowania wpływają zmiany parametrów m_x i m_y .

Tabela 2.1

Nazwa obrazu	t [min.]	l
2a	10,3	10
2b	9,28	9
2c	21,6	14
2d	20	13

Tabela 2.1 zawiera informacje o procesie klasteryzacji obrazów 2a, 2b, 2c, 2d. W procesie podziału obrazów 2a, 2b, 2d użyto 16 klastery. W przypadku obrazu 2c użyto 25 klastery.



Rys. 3. Trzecia sekwencja obrazów a, b, c, d i rozkład klastery w nich zawartych e, f, g, h

Fig. 3. Third sequence of images a, b, c, d and distribution of clusters e, f, g, h

Kolejny przykład (rys. 3) dotyczy sytuacji, w której na obrazie pojawia się zakłócenie w postaci obiektu o niewielkiej deformacji kształtu. We wszystkich przypadkach nastąpiło pasowanie trójkąta z trójkątem o zdeformowanym kształcie. Wynika to z jego bliższego położenia względem trójkąta z obrazu 3a.

Tabela 3

Spasowane klasterzy	Δm_x	Δm_y	$\Delta m_z / m_{za}$	$\Delta m_m / m_{ma}$	$\Delta m_s / m_{sa}$	D
3a[1]-3b[1]	132,36	-142,33	0	0,12	0	640007,08
3a[1]-3b[2]	185,84	-204,22	0	0,01	0	1255228,35
3a[1]-3c[1]	4,26	-105,83	0	0,19	-0,04	242664,15
3a[1]-3c[2]	9,54	-202,82	0	-0,01	-0,07	599944,57
3a[1]-3d[1]	11,82	-11,09	0	0,2	0,04	237459,24
3a[1]-3d[2]	184,82	-204,96	0	0,02	-0,67	51003908,79

Tabela 3 przedstawia wyniki pasowania klasterów. We wszystkich przypadkach nastąpiło błędne pasowanie. Dla klastera 3a[1] z obrazu 3a algorytm zawsze odnajdywał odpowiednik na obrazach 3c i 3d w klasterze przedstawiającym zdeformowany trójkąt (oznaczony numerem 1). Następnym najbliższym klasterem we wszystkich przypadkach okazywał się trójkąt.

Dane z tabeli 3 pokazują, że poza przemieszczeniem w znacznym stopniu zmienia się również parametr $\Delta m_m / m_{ma}$. Zmiany jego wartości określają poprawne pasowanie, jednakże są one zbyt małe, by kompensować duże zmiany położenia obiektu. W ostatnim przypadku na niekorzyść prawidłowego pasowania wpływa również duża wielkość:

$$\Delta m_s / m_{sa} = -0,67$$

W tabeli 3.1 zamieszczono dane dotyczące klasteryzacji obrazów 3a, 3b, 3c, 3d. Dla wszystkich obrazów do podziału użyto 16 klasterów.

Tabela 3.1

Nazwa obrazu	t [min.]	l
3a	10,5	11
3b	11,5	12
3c	11,5	12
3d	10,5	11

W przypadku obrazów 3a, 3b, 3c, 3d oraz 2d wystąpiły kłopoty z poprawnym pasowaniem obiektów sceny. W obu przypadkach błędne wyniki pasowania są spowodowane dużymi zmianami parametrów m_x i m_y .

4. Modyfikacja metody

Na podstawie wyników przeprowadzonych badań stwierdzono, że dla poprawy efektywności metody należy wyeliminować niekorzystny wpływ tych parametrów (m_x, m_y), które uniemożliwiają poprawne spasowanie odpowiadających sobie klasterów w przypadku, kiedy występuje pomiędzy nimi duże przemieszczenie. W celu wymuszenia poprawnego pasowania figur eksperyment powtórzono modyfikując proces optymalizacji wag. Momenty środków ciężkości m_x i m_y zastąpiono deskryptorami kształtu. Deskryptory te są długościami osi elipsy równoważnej, tj. takiej elipsy, której momenty bezwładności są identyczne z momentami danego klastera [5],[6]. Wyrażają się one następującym wzorem:

$$\lambda_{1,2} = \frac{\sigma_x + \sigma_y}{2} \pm \frac{\sqrt{(\sigma_x + \sigma_y)^2 + 4\sigma_{xy}^2}}{2},$$

gdzie:

$$\sigma_x = \frac{\sum_{x,y \in k} (x - \bar{x})^2}{N}, \quad \sigma_y = \frac{\sum_{x,y \in k} (y - \bar{y})^2}{N}, \quad \sigma_{xy} = \frac{\sum_{x,y \in k} (x - \bar{x})(y - \bar{y})}{N},$$

$$\bar{x} = \frac{\sum_{x,y \in k} x}{N}, \quad \bar{y} = \frac{\sum_{x,y \in k} y}{N}$$

Na podstawie powyższych zależności oblicza się obrót obiektu korzystając z następującego wzoru:

$$\theta = \theta_1 - \theta_2,$$

gdzie θ jest kątem obrotu, θ_1 i θ_2 są orientacjami (kątami nachylenia długich osi tych elips do osi OX na obrazach) elips równoważnych klastera w obu położeniach, wyznaczanych z następującego wzoru:

$$\theta = \frac{1}{2} \operatorname{arccctg} \left(\frac{\sigma_y - \sigma_x}{2\sigma_{xy}} \right)$$

Należy przy tym pamiętać, że kąt obrotu obiektu wykrywany tą metodą jest ograniczony do przedziału $\left[-\frac{\pi}{2}, \frac{\pi}{2} \right]$.

Powtarzając proces pasowania dla sekwencji 3 uzyskano oczekiwane wyniki. Ulepszona metoda sprawdziła się również w przypadku pasowania obrazów 2a i 2d.

Uwaga: Rząd średniej złożoności obliczeniowej całego algorytmu w przypadku zastosowania jego zmodyfikowanej wersji wynosi:

$$O(T) = O(n^2),$$

gdzie: n - liczba klasterów.

Tabela 4

Spasowane klasterzy	$\Delta\lambda_1$	$\Delta\lambda_2$	$\Delta m_z/m_{za}$	$\Delta m_m/m_{ma}$	$\Delta m_s/m_{sa}$	D
3a[1]-3b[2]	0,1	-0,1	0	0,01	0	41521,41
3a[1]-3b[1]	1,44	0,26	0	0,12	0	55774,77
3a[1]-3c[2]	0,04	-0,09	0	-0,01	-0,07	889,91
3a[1]-3c[1]	1,43	1,28	0	0,19	-0,04	203103,93
3a[1]-3d[2]	0,21	0,03	0	0,02	0,67	2731,53
3a[1]-3d[1]	1,36	1,25	0	0,2	-0,06	359552,71

OZNACZENIA:

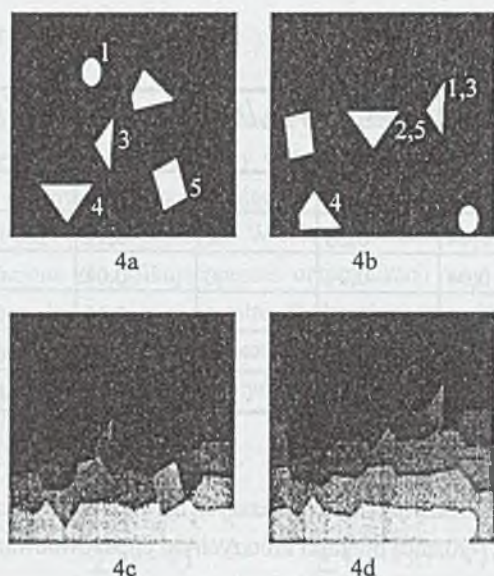
$\Delta\lambda_1 = \lambda'_1 - \lambda_1$ - różnica długości dłuższych osi elips równoważnych,

$\Delta\lambda_2 = \lambda'_2 - \lambda_2$ - różnica długości krótszych osi elips równoważnych.

Tabela 4 przedstawia wyniki pasowania klasterów przy zastosowaniu zmodyfikowanej metody. We wszystkich przypadkach nastąpiło poprawne pasowanie. Dla klastera 3a[1] z obrazu 3a algorytm zawsze odnajdywał odpowiednik na obrazach 3b, 3c i 3d w klasterze przedstawiającym trójkąt. Tym razem następnym najbliższym klasterem we wszystkich przypadkach okazał się ten, który reprezentuje zdeformowany trójkąt.

Dane w tabeli 4 ilustrują korzystny wpływ parametrów λ_1 i λ_2 na poprawny wynik pasowania obiektów.

W celu porównania metody i jej zmodyfikowanej wersji pasowaniu poddano dwa obrazy o większej złożoności (rys. 4). Znajduje się na nich pięć figur o różnym kształcie. Stosując metodę opisaną w [1], tylko figura przedstawiająca mały trójkąt została spasowana ze swoim odpowiednikiem. Jednak przy zastosowaniu metody zmodyfikowanej dla wszystkich figur udało się uzyskać pozytywny wynik. Używając metody opisanej w [1] spasowano pary figur o tych samych numerach. Jak pokazano, mały trójkąt nr 3 z obrazu 4a został prawidłowo spasowany ze swoim odpowiednikiem na obrazie 4b (trójkąt nr 1), ale elipsa nr 1 na obrazie 4a również została spasowana z tym samym trójkątem (stąd dwa numery 1,3). Podobnie dzieje się z trójkątem 2,5 z obrazu 4b. Natomiast prostokąt oraz elipsa z obrazu 4b nie posiadają odpowiedników na obrazie 4a.



Rys. 4. Czwarta sekwencja obrazów

Fig. 4. Fourth sequence of images

5. Wnioski

Porównując sposób, w jaki przebiega proces pasowania metodą opisaną w [1] i jej zmodyfikowaną wersją zaproponowaną w niniejszym opracowaniu, możemy zauważyć że:

1. Stosowanie metody uwzględniającej przesunięcie klastra w osi X i Y jest skuteczniejsze w przypadku analizy scen perspektywicznych, w których obiekty ruchome są bryłami trójwymiarowymi, a ich przemieszczenie jest niewielkie. Zauważono również, że można odwrócić kolejność pasowania niektórych obrazów (tzn. zamienić miejscami kolejne ramki, np. 2a z 2d), jeśli proces pasowania przebiegł niewłaściwie. Podobnie jak w przypadku pasowania obrazu 2a z 2d, zamiana miejscami kolejnych obrazów może spowodować poprawne pasowanie.
2. Skuteczność metody zmodyfikowanej jest wysoka dla sekwencji obrazów zawierających elementy, których kształt i rozkład jasności w sekwencji w bardzo małym stopniu ulegają zmianie. Zakres zmiany położenia nie jest tutaj istotny. Metoda sprawdza się lepiej podczas analizy scen naturalnych o niewielkiej głębi. Efektywność pasowania może się zwiększyć, jeżeli możliwe jest przeprowadzenie wstępnej segmentacji (np. binaryzacja, progowanie, operacje morfologiczne).

Poprawność procesu pasowania w obu przypadkach zależy od właściwego przebiegu klasteryzacji obrazów. Podstawową zasadą poprawnej klasteryzacji jest ustalenie rozmiarów² klastery początkowych tak, by były zbliżone do rozmiarów obiektów. Niestety, nie wystarcza to do właściwej klasteryzacji. Problem niewłaściwego podziału występuje wtedy, gdy obiekt o jednolitej jasności reprezentowany jest przez więcej niż jeden klastery. Dzieje się tak co najmniej z trzech powodów. Pierwszy wynika ze złożoności kształtu obiektów (np. gdy obiekt jest podłużny lub jego kontur ma dużo wypukłości i wklęsłości). Drugi z dużego zróżnicowania wielkości znajdujących się na scenie obiektów, dla których nie da się dobrać odpowiedniej wielkości klastery początkowych. Trzeci problem pojawia się, gdy obiekt znajduje się na granicy klastery początkowych. Istnieje wtedy duże prawdopodobieństwo reprezentacji obiektu przez dwa lub więcej klastery.

Zmodyfikowana metoda pozwala obliczać kąt obrotu obiektu w zakresie $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.

Ograniczenie to wynika z efektu braku rozróżnialności większych obrotów, gdyż osie elipsy równoważnej, opisujące jednakowe pod względem kształtu klastery, powracają do tego samego położenia po wykonaniu obrotu π lub $-\pi$. Rozwiązaniem tego problemu może być np. znalezienie kąta między odcinkiem łączącym środek ciężkości klastera i najbliższy mu punkt leżący na jego konturze a osią X na obrazie, co będzie celem następnych modyfikacji algorytmu.

LITERATURA

1. Dane P. Kottke, Ying Sun: *Motion Estimation via Cluster Matching*, IEEE Trans. PAMI, vol. 16, no. 11, pp. 1128-1132, november 1994.
2. G. Shunck: *Image Flow Segmentation and Estimation by Constraint Line Clustering*, IEEE Trans. PAMI, vol. 11, no. 10 pp. 1010-1027, 1989.
3. S. Beauchemin, J. L. Barron: *The Computation of Optical Flow*, ACM Computing Surveys, vol. 27, no. 3, pp 434-467, 1995.
4. Xiaoping Hu, Narendra Ahuja: *Feature Extraction and Matching as Signal Detection*, Inte. Journal of PAMI vol. 8, no. 6, pp. 1343-1379, 1994.
5. Willima K. Pratt: *Digital Image Processing*, Johs Wiley & Sons, Inc. pp. 629-646, USA, 1991.
6. Simon X. Liao, Mirosław Pawlak: *On Image Analysis by Moments*, IEEE Trans. PAMI, vol. 18, no. 3, 1996.

² Początkowo obraz zostaje podzielony na K kwadratowych regionów. Liczba klastery jest odwrotnie proporcjonalna do dół ich powierzchni.

7. Alessandro Verri, Tomaso Poggio: *Motion Field and Optical Flow: Qualitative Properties*, IEEE Trans. PAMI, vol. 11, no. 5, pp. 490-498, 1989.

Recenzent: Dr hab.inż. Konrad Wojciechowski
Prof. Politechniki Śląskiej

Wpłynęło do Redakcji 1.10.1997 r.

Abstract

The detection algorithm presented in the article differs from typical methods of a motion detection. The matching refers to regions of images (objects) which are coherent with regard to their intensity. They are obtained in the clustering process (hence these regions are called clusters). The moving objects are represented by shifted clusters which have similar intensity and shape features.

There are at least two reasons that make this method interesting:

- generation of a more regular optical flow than in case of other methods;
- detection of a motion on a large scale changes of an object location in the image as well as changes of intensity and the shape deformation of an object.