AUTOMATYKA • 2008 • Tom 12 • Zeszyt 3

Przemysław Korohoda*

Porównanie dwóch metod trójwymiarowego modelowania obrazu za pomocą szeregu Taylora drugiego rzędu**

1. Wprowadzenie

Opublikowany w roku 2004 algorytm SIFT (*Scale-Invariant Feature Transform*) [7] stanowi interesującą i często cytowaną propozycję metody porównującej dwa obrazy tej samej sceny, różniące się skalą lub lokalizacją kamery względem sceny. Algorytm ten bazuje na wyznaczeniu w porównywanych obrazach punktów charakterystycznych [6]. Po ich wstępnym wyborze realizowana jest selekcja, wykorzystująca rozwinięcie rozważanego fragmentu obrazu (bloku 3×3 piksele), otaczającego punkt charakterystyczny, w szereg Taylora. Szereg ten jest ograniczony do pochodnych drugiego rzędu, co oznacza, iż badany fragment jest aproksymowany paraboloidą. Podobną aproksymację stosuje się również w trójwymiarowym modelowaniu treści obrazu cyfrowego [11]

Konieczne w trakcie obliczeń wyznaczanie hesjanu nie jest zazwyczaj opisywane w podręcznikach przetwarzania obrazów [2, 4, 9]. Z kolei zaawansowane opisy algorytmów pomijają istotne szczegóły obliczeniowe, odwołując się do ogólnej wiedzy w tej dziedzinie lub do wcześniejszych publikacji. Przykładowo w [7] wspomniano, iż konieczne obliczenia można realizować w sposób prosty, odsyłając czytelnika do [3], gdzie jednak całość zagadnienia została przedstawiona jednym ogólnym zdaniem, już bez dalszych odnośników. Z kolei w [11] opisano metodykę postępowania za pomocą ogólnych równań macierzowych, nierozważając możliwości szybkich obliczeń za pomocą filtrów splotowych.

Dokładniejsze zbadanie zagadnienia wyznaczania hesjanu dla bloków 3×3 pikseli wykazało, iż do rozwiązania postawionego problemu można zastosować różne metody, których wyniki w pewnych przypadkach mogą się zasadniczo różnić. Podane argumenty skłoniły autora do sporządzenia opracowania porządkującego wyniki własnych obserwacji z odpowiednim wyborem niezbędnych informacji o charakterze tutorialnym. Wszystkie wyprowadzenia zostały przeprowadzone przez autora niezależnie, by mieć pewność, iż

* Katedra Elektroniki, Akademia Górniczo-Hutnicza w Krakowie
** Praca realizowana w ramach prac statutowych AGH, nr 11.11.120.766

przedstawiony opis jest kompletny i by mógł stanowić użyteczną pomoc w praktycznej realizacji algorytmów przetwarzania obrazów, korzystających z aproksymacji bloków 3×3 za pomocą paraboloidy.

Warto wspomnieć o jeszcze jednym możliwym, a niestosowanym powszechnie, wykorzystaniu interpolacji (aproksymacji) paraboloidalnej. Interpolacja międzypikselowa obrazów cyfrowych, np. na potrzeby zmiany rozdzielczości, jest zazwyczaj realizowana z wykorzystaniem trzech najbardziej popularnych technik interpolacyjnych, czyli metodami:

a) najbliższego sąsiada (*nearest neighbour*),
b) interpolacji dwuliniowej (*bilinear*),
c) interpolacji dwusześciennej (*bicubic*).

Opisana powyżej interpolacja paraboloidą może być traktowana jako **dwukwadratowa**, a zatem stanowi rozwiązanie pośrednie między metodami b) i c).

Zawartość pozostałej części artykułu jest następująca. W rozdziale drugim przedstawiono i porównano różne formy zapisu szeregu Taylora. Kolejny rozdział zawiera wyprowadzenie dwóch metod wyznaczania współczynników szeregu Taylora dla bloku 3×3 piksele wraz z przykładem wykazującym istotne różnice w wynikach stosowania obu metod. W rozdziale czwartym wyznaczono maski filtrów splotowych, przy wykorzystaniu których obie metody można realizować za pomocą łatwych w stosowaniu filtrów splotowych. Rozdział piąty zawiera wyprowadzenie proponowanego przez autora współczynnika liczbowego do lokalizacji grzbietu w obrazie, stanowiącego rozwiązanie alternatywne wobec proponowanej w [7] proporcji wartości własnych hesjanu. W rozdziale szóstym przedstawiono przykład ilustrujący wcześniejsze rozważania. Uwagi podsumowujące zawiera rozdział siódmy.

2. Dwuwymiarowy szereg Taylora drugiego rzędu

Przytaczany w literaturze [3, 7] szereg Taylora można także nazywać szeregiem Maclaurina, ponieważ zakłada przy rozważaniu każdego bloku 3×3 przeniesienie środka układu współrzędnych do punktu centralnego danego bloku. Ze względu na różne, stosowane w literaturze, postacie zapisu, poniżej zostaną zebrane najważniejsze z nich, z podkreśleniem wzajemnych zależności, umożliwiających odpowiednie przeliczenia.

Przyjmijmy, iż osie dwuwymiarowego układu współrzędnych x oraz y są ciągłe, przy czym w przetwarzaniu obrazów zazwyczaj zakłada się, że oś x jest pionowa, skierowana w dół, a oś y jest pozioma, skierowana w prawo. W takim przypadku szereg aproksymujący funkcję f w otoczeniu punktu o współrzędnych (0, 0) może być wyrażony następującym wzorem

$$\begin{aligned} f(x,y) &= f(0,0) + f_x(0,0)\cdot x + f_y(0,0)\cdot y + \\ &+ \frac{1}{2}\cdot f_{xx}(0,0)\cdot x^2 + \frac{1}{2}\cdot f_{yy}(0,0)\cdot y^2 + f_{xy}(0,0)\cdot x\cdot y \end{aligned} \tag{1}$$

gdzie przez f_x, f_y, f_{xx}, f_{yy}, f_{xy} oznaczono odpowiednie pochodne cząstkowe pierwszego i drugiego rzędu. We wzorze (1) przyjęto ponadto, że: $f_{xy} = f_{yx}$.

Tę samą zależność można także zapisać w postaci wielomianu kwadratowego zmiennych x oraz y:

$$f(x,y) = A \cdot x^2 + B \cdot y^2 + C \cdot x \cdot y + D \cdot x + E \cdot y + F \tag{2}$$

przy czym wzajemna zależność między współczynnikami (1) i (2) wynika wprost z porównania wielomianów

$$\begin{aligned} &A = \frac{1}{2} \cdot f_{xx}(0,0); \quad B = \frac{1}{2} \cdot f_{yy}(0,0); \quad C = f_{xy}(0,0); \\ &D = f_x(0,0); \quad E = f_y(0,0); \quad F = f(0,0) \end{aligned} \tag{3}$$

Zapis (1) można również przedstawić, eksponując wektorową interpretację pary współrzędnych (x, y)

$$f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = f\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) + \begin{bmatrix} f_x & f_y \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \frac{1}{2} \cdot \begin{bmatrix} x & y \end{bmatrix} \cdot \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \tag{4}$$

W (4) nie zaznaczono (przyjmując to jako oczywiste), iż pochodne cząstkowe liczone są w środku układu współrzędnych. Forma wektorowa stwarza możliwość zdefiniowania wektora gradientu, **g**, oraz macierzy Hessego, **H**, zwanej hesjanem. W poniższych równaniach przedstawiono definicje odpowiednich wektorów i macierzy

$$\bar{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}; \quad \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}; \quad \mathbf{g} = \begin{bmatrix} f_x \\ f_y \end{bmatrix}; \quad \mathbf{H} = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \tag{5}$$

Korzystając z tych definicji, można równanie (1) przepisać do postaci macierzowej

$$f(\mathbf{x}) = f(\bar{0}) + \mathbf{g}^T \cdot \mathbf{x} + \frac{1}{2} \cdot \mathbf{x}^T \cdot \mathbf{H} \cdot \mathbf{x} \tag{6}$$

Warto zaznaczyć, iż równanie (6) może być stosowane również dla więcej niż dwóch wymiarów.

Stosowanie opisanej metody aproksymacji polega na tym, iż współczynniki rozwinięcia wyznaczanie są dla każdego rozważanego piksela, którego położenie przyjmuje się jako (0, 0), na podstawie bloku 3×3 pikseli, którego centralnym punktem jest rozważany piksel. Wyznaczone współczynniki mogą być następnie wykorzystywane do dalszych obliczeń – jako element algorytmu SIFT [7], jako niezależne narzędzie do wyznaczania punktów charakterystycznych [8] lub w celu modelowania kształtu w trzech wymiarach [11].

3. Wyznaczanie współczynników szeregu Taylora

Jako założenie wstępne przyjmujemy, iż znane są wartości 9 punktów bloku 3×3. Przyjmując, iż odległości międzypikselowe w pionie i w poziomie wynoszą 1, otrzymuje

się, że wartości współrzędnych znanych punktów należą do zbioru {-1, 0, +1}. Jeżeli blok 3×3 pikseli potraktujemy jak macierz **B**, to poszczególne piksele można oznaczyć następująco

$$\mathbf{B} = \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix} \tag{7}$$

Piksel *e* jest zatem pikselem centralnym bloku **B**.

3.1. Metoda z uśrednianiem

Dla każdego rozważanego piksela należy wyznaczyć 6 współczynników szeregu (2), które można zapisać w postaci wektorowej

$$\mathbf{c} = [A, B, C, D, E, F]^T \tag{8}$$

Blok 3×3 dostarcza 9 wartości, które umożliwiają utworzenie 9 równań. Oznacza to, iż liczba równań przekracza liczbę niewiadomych. Jednym z możliwych rozwiązań tego problemu jest czterokrotne wyliczenie 6 współczynników – za każdym razem bazując na innym punkcie narożnym bloku 3×3 – i arytmetycznym uśrednieniu otrzymanych wyników. Jeżeli zapiszemy wartości pikseli bloku **B** w postaci wektorowej

$$\mathbf{p} = [a, b, c, d, e, f, g, h, i]^T \tag{9}$$

to wybór kolejnych narożników (punkty *a*, *c*, *g*, *i*) prowadzi do czterech 6-elementowych wektorów reprezentujących wartości znane:

$$\begin{aligned} \mathbf{p}_a &= [a, b, d, e, f, h]^T; \quad \mathbf{p}_c = [b, c, d, e, f, h]^T \\ \mathbf{p}_g &= [b, d, e, f, g, h]^T; \quad \mathbf{p}_i = [b, d, e, f, h, i]^T \end{aligned} \tag{10}$$

Dla każdego z powyższych wektorów można ułożyć odpowiednie równanie

$$\mathbf{X}_k \cdot \mathbf{c}_k = \mathbf{p}_k : \quad k \in \{a, c, g, i\} \tag{11}$$

gdzie wiersze kwadratowej macierzy $\mathbf{X}_k$ zawierają kolejno, odpowiednie dla każdego piksela, wartości: x^2, y^2, $x{\cdot}y$, x, y oraz e (czyli wartość piksela centralnego w danym bloku 3×3). Rozwiązanie równań (11) jest następujące

$$\mathbf{c}_k = \mathbf{X}_k^{-1} \cdot \mathbf{p}_k \tag{12}$$

Ostateczne wartości współczynników równania (2) otrzymujemy uśredniając arytmetycznie wyniki (12)

$$\mathbf{c} = \frac{\mathbf{c}_a + \mathbf{c}_c + \mathbf{c}_g + \mathbf{c}_i}{4} \tag{13}$$

Należy podkreślić, że metoda ta gwarantuje, iż wartość piksela *e* po zastosowaniu aproksymacji (2) będzie zachowana.

3.2. Metoda z macierzą pseudoodwrotną

Jak już stwierdzono, jeżeli wszystkie wartości pikseli bloku 3×3 potraktujemy jako niezależne wartości dane, to ich liczba przekracza liczbę szukanych wartości zawartych w wektorze **c** (8). Prowadzi to do następującego równania macierzowego

$$\mathbf{X} \cdot \mathbf{c} = \mathbf{p} + \mathbf{e} \tag{14}$$

gdzie **e** jest wektorem błędu w odniesieniu do wartości zawartych w wektorze **p**. W tym przypadku kolejne wiersze macierzy **X** (o wymiarach 6×9) zawierają analogiczne elementy jak w przypadku macierzy $\mathbf{X}_k$ w równaniach (11). Oznaczmy jako *SE* (*square error*) wartość sumy kwadratów błędów zawartych w wektorze **e**. Po przekształceniach mamy

$$SE = \mathbf{e}^T \cdot \mathbf{e} = (\mathbf{X} \cdot \mathbf{c} - \mathbf{p})^T \cdot (\mathbf{X} \cdot \mathbf{c} - \mathbf{p}) = \mathbf{c}^T \cdot \mathbf{X}^T \cdot \mathbf{X} \cdot \mathbf{c} - 2 \cdot \mathbf{p}^T \cdot \mathbf{X} \cdot \mathbf{c} + \mathbf{p}^T \cdot \mathbf{p} \tag{15}$$

Jest to forma kwadratowa ze względu na wektor **c**, reprezentujący argument, zatem minimum wartości (15) może być wyznaczone przez przyrównanie pochodnej (liczonej wektorowo po **c**) z *SE* do zera

$$\frac{\delta}{\delta \mathbf{c}} SE = 2 \cdot \mathbf{c}^T \cdot \mathbf{X}^T \cdot \mathbf{X} - 2 \cdot \mathbf{p}^T \cdot \mathbf{X} \quad \Rightarrow \quad \mathbf{X}^T \cdot \mathbf{X} \cdot \mathbf{c} - \mathbf{X}^T \cdot \mathbf{p} = 0 \tag{16}$$

Przekształcając (16), otrzymujemy zatem przepis na wektor **c** minimalizujący wartość *SE*

$$\mathbf{c} = \left(\mathbf{X}^T \cdot \mathbf{X}\right)^{-1} \cdot \mathbf{X}^T \cdot \mathbf{p} \tag{17}$$

Macierz

$$\mathbf{A} = (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T \tag{18}$$

zwana jest macierzą pseudoodwrotną do niekwadratowej macierzy **X**.

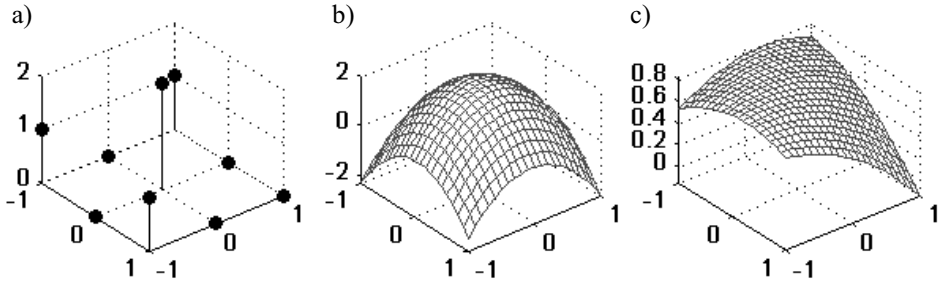
Równanie macierzowe (17) opisuje zatem drugą metodę wyznaczania współczynników rozwinięcia w szereg (2) dla danego bloku 3×3.

W przypadku tej metody wartość piksela centralnego *e* nie jest zazwyczaj zachowana, natomiast aproksymacja prowadzi do minimalizacji błędu *SE*, liczonego po wszystkich dziewięciu pikselach bloku.

Można zatem stwierdzić, iż obie opisane metody dają identyczne wyniki, gdy wartości pikseli od początku leżą na powierzchni paraboloidy. Wtedy szereg (2) opisuje funkcję interpolującą między pikselami. Funkcja staje się aproksymującą, gdy wartości pikseli nie da się umieścić na powierzchni paraboloidy, wtedy też ujawniają się różnice między opisanymi metodami.

3.3. Przykład

Przyjmijmy blok 3×3 pikseli o wartościach, które można następująco zapisać w postaci wektora **p** (patrz (7) i (9)): $\mathbf{p} = [1\ 0\ 1\ 0\ 2\ 0\ 1\ 0\ 0]^T$. Zastosowanie opisanych wyżej metod prowadzi do współczynników **c**, definiujących funkcje (2), których wykresy przedstawiono na rysunku 1.



Rys. 1. Wynik aproksymacji za pomocą paraboloidy przykładowego bloku 3×3 z wykorzystaniem dwóch rozważanych metod: a) rozmieszczenie zadanych punktów; b) metoda z uśrednieniem; c) metoda z macierzą pseudoodwrotną

Lokalną aproksymację funkcji obrazu z wykorzystaniem hesjanu stosuje się np. w algorytmie SIFT w celu stwierdzenia obecności **grzbietu** [7]. Nawet wstępne wizualne porównanie wykazuje, iż metoda uśredniająca obecności grzbietu nie wykaże, natomiast metoda z macierzą pseudoodwrotną – tak. Porównanie może być przeprowadzone w sposób formalny. W [7] w celu wykrycia obecności grzbietu korzysta się z parametru liczbowego r wyrażającego proporcję wartości własnych hesjanu **H** (9)

$$r = \frac{\lambda_1}{\lambda_2} \tag{19}$$

Parametr (19) określa relację między dwiema krzywiznami głównymi. Na obecność grzbietu wskazuje według [7] wartości r większa od zadanego progu. Przyjmuje się, iż wartości własne są uporządkowane w module: $|\lambda_1| = |\lambda_2|$, a pod uwagę bierze się tylko te piksele, dla których wartości własne są tego samego znaku. Różnica znaku wartości własnych świadczy o tym, że mamy do czynienia z tzw. punktem siodłowym. Zerowanie się mniejszej wartości własnej musi być traktowane jako przypadek szczególny (19). Gdy obie krzywizny główne są zbliżone, czyli dla $r \approx 1$, wówczas przyjmuje się, iż grzbiet nie występuje. Obecność grzbietu ujawnia się w postaci różnicy między krzywiznami, których miarę stanowią wartości własne.

W przypadku przedstawionym na rysunku 1, dla a) r = 1,13(3), dla b) r = 7, co jednoznacznie wskazuje na poważną różnicę między metodami. Niekiedy w roli parametru stosuje się wyznacznik hesjanu [8], który dla powyższego przykładu przyjmuje następujące wartości: a) 15,94 >> 0, b) 0,05, co ponownie wskazuje na poważne różnice, w zależności od zastosowanej metody.

4. Szybkie wyznaczanie współczynników rozwinięcia

W wyrażeniu na rozważany szereg Taylora występują pochodne kierunkowe oraz pochodne drugiego rzędu. Dla wszystkich tych wielkości można znaleźć w literaturze propozycje filtrów splotowych zdefiniowanych w postaci odpowiednich masek (macierze 3×3) [2, 4, 9], których zastosowanie ma zastępować wyliczenie odpowiedniej pochodnej. Stosowanie tych filtrów jest zazwyczaj opisywane jako dyskretna aproksymacja odpowiednich pochodnych ciągłych, ponieważ większość z tych masek jest wyprowadzania na podstawie odpowiednich równań różnicowych. Dla opisanych wyżej metod można wyprowadzić analogiczne maski, korzystając z równań (12) (13) oraz (17). Wszystkie macierze: $\mathbf{X}$ oraz $\mathbf{X}_k$ są ustalone z góry, zatem rozpisując równania (13) i (17) dla kolejnych elementów $\mathbf{c}$ (8) i korzystając z zależności (3), można określić, z jakimi współczynnikami wagowymi odpowiednie wartości pikseli wchodzą do rozwiązania dla wartości każdego elementu wektora $\mathbf{c}$. Tak określone współczynniki wagowe można przedstawić w postaci masek filtru splotowego. Dla metody z uśrednianiem otrzymuje się w ten sposób następujący zbiór masek 3×3, zapisanych w postaci macierzy:

$$\mathbf{M}_{xx} = \begin{bmatrix} 0 & +1 & 0 \\ 0 & -2 & 0 \\ 0 & +1 & 0 \end{bmatrix}; \quad \mathbf{M}_{yy} = \begin{bmatrix} 0 & 0 & 0 \\ +1 & -2 & +1 \\ 0 & 0 & 0 \end{bmatrix}; \quad \mathbf{M}_{xy} = \frac{1}{4}\cdot\begin{bmatrix} +1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & +1 \end{bmatrix}$$
$$\mathbf{M}_{x} = \frac{1}{2}\cdot\begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & +1 & 0 \end{bmatrix}; \quad \mathbf{M}_{y} = \frac{1}{2}\cdot\begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & +1 \\ 0 & 0 & 0 \end{bmatrix} \tag{20}$$

Dla metody z macierzą pseudoodwrotną odpowiednie maski są następujące:

$$\mathbf{W}_{xx} = \frac{1}{3}\cdot\begin{bmatrix} +1 & +1 & +1 \\ -2 & -2 & -2 \\ +1 & +1 & +1 \end{bmatrix}; \quad \mathbf{W}_{yy} = \frac{1}{3}\cdot\begin{bmatrix} +1 & -2 & +1 \\ +1 & -2 & +1 \\ +1 & -2 & +1 \end{bmatrix}; \quad \mathbf{W}_{xy} = \frac{1}{4}\cdot\begin{bmatrix} +1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & +1 \end{bmatrix}$$
$$\mathbf{W}_{x} = \frac{1}{6}\cdot\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ +1 & +1 & +1 \end{bmatrix}; \quad \mathbf{W}_{y} = \frac{1}{6}\cdot\begin{bmatrix} -1 & 0 & +1 \\ -1 & 0 & +1 \\ -1 & 0 & +1 \end{bmatrix}; \quad \mathbf{W} = \frac{1}{9}\cdot\begin{bmatrix} -1 & +2 & -1 \\ +2 & +5 & +2 \\ -1 & +2 & -1 \end{bmatrix} \tag{21}$$

Warto zwrócić uwagę na fakt, iż w obu metodach maski dla pochodnych mieszanych są takie same.

5. Alternatywny współczynnik do wykrywania grzbietu

Zbadanie sposobów wyznaczania hesjanów **H** (5) dla dwóch rozważanych metod wykazało, iż macierze te różnią się jedynie o macierz jednostkową **E** mnożoną przez współczynnik γ

$$\mathbf{H}_2 = \mathbf{H}_1 - \gamma \cdot \mathbf{E} \quad : \quad \gamma = \frac{1}{3} \cdot \left[2 \cdot (b + d + f + h) - (a + c + g + i) - 4 \cdot e \right] \tag{22}$$

Wyznaczanie współczynnika γ dla każdego piksela może być zrealizowane przez filtrację splotową z wykorzystaniem następującej maski filtru

$$\mathbf{W}_\gamma = \frac{1}{3} \cdot \begin{bmatrix} -1 & 2 & -1 \\ 2 & -4 & 2 \\ -1 & 2 & -1 \end{bmatrix} \tag{23}$$

Jeżeli przyjmiemy, iż układ ortonormalnych wektorów własnych hesjanu $\mathbf{H}_1$ stanowi kolejne kolumny macierzy **V**, natomiast $\mathbf{L}_1$ jest macierzą diagonalną, zawierającą na przekątnej kolejne wartości własne macierzy $\mathbf{H}_1$, to prawdziwa jest poniższa sekwencja przekształceń [1, 5, 10]

$$\begin{aligned} \mathbf{V}^T \cdot \mathbf{H}_2 \cdot \mathbf{V} &= \mathbf{V}^T \cdot (\mathbf{H}_1 - \gamma \cdot \mathbf{E}) \cdot \mathbf{V} = \mathbf{V}^T \cdot \mathbf{H}_1 \cdot \mathbf{V} - \gamma \cdot \mathbf{V}^T \cdot \mathbf{E} \cdot \mathbf{V} = \\ &= \mathbf{V}^T \cdot \mathbf{H}_1 \cdot \mathbf{V} - \gamma \cdot \mathbf{E} = \mathbf{L}_1 - \gamma \cdot \mathbf{E} \end{aligned} \tag{24}$$

Otrzymana na końcu wyprowadzenia (24) macierz jest diagonalna, zatem macierz $\mathbf{H}_2$ posiada identyczny zbiór wektorów własnych co $\mathbf{H}_1$. Otrzymana w (24) macierz diagonalna zawiera na przekątnej wartości własne macierzy $\mathbf{H}_2$, czyli

$$\mathbf{L}_2 = \mathbf{L}_1 - \gamma \cdot \mathbf{E} \tag{25}$$

Z macierzowego równania (25) wynika, iż prawdziwe są zależności skalarne:

$$\lambda_{21} = \lambda_{11} - \gamma; \quad \lambda_{22} = \lambda_{12} - \gamma \tag{26}$$

gdzie:

$\lambda_{11}, \lambda_{12}$ – wartości własne dla $\mathbf{H}_1$,
$\lambda_{21}, \lambda_{22}$ – wartości własne $\mathbf{H}_2$.

Odejmując (26) stronami, otrzymujemy, iż

$$\lambda_{11} - \lambda_{12} = \lambda_{21} - \lambda_{22} \quad \Rightarrow \quad |\lambda_{11} - \lambda_{12}| = |\lambda_{21} - \lambda_{22}| \tag{27}$$

Oznacza, to że moduły różnic wartości własnych dla hesjanów w przypadku obu porównywanych metod są takie same. Na podstawie powyższego wyprowadzenia proponuje się zatem nowy parametr liczbowy do badania lokalnego kształtu funkcji obrazu, zastępujący parametr (19), w następującej postaci

$$R = |\lambda_1 - \lambda_2| \tag{28}$$

Dodatkową zaletą parametru R jest łatwość jego wyznaczania bez konieczności wyliczania wartości własnych – podobnie, jak w przypadku parametru r. Suma wartości własnych jest równa śladowi macierzy (operator: tr), natomiast ich iloczyn równy jest wyznacznikowi (operator: det) [1, 5, 10]:

$$\lambda_1 + \lambda_2 = \mathrm{tr}(\mathbf{H}) = f_{xx} + f_{yy} \tag{29}$$

$$\lambda_1 \cdot \lambda_2 = \det(\mathbf{H}) = f_{xx} \cdot f_{yy} - f_{xy}{}^2 \tag{30}$$

Bazując na następującej zależności algebraicznej (dla dowolnych liczb a oraz b)

$$(a-b)^2 = (a+b)^2 - 4 \cdot a \cdot b \tag{31}$$

otrzymujemy przepis na wyliczenie kwadratu poszukiwanego wyrażenia (28)

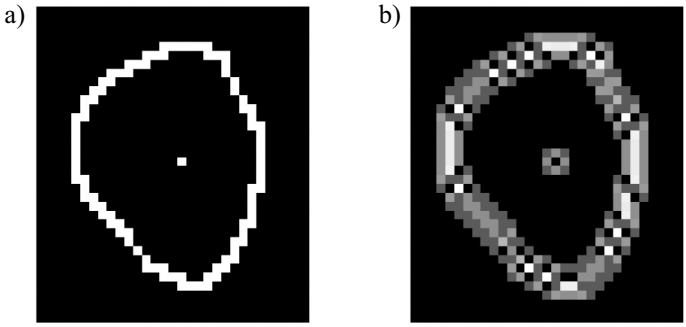
$$(\lambda_1 - \lambda_2)^2 = f_{xx}{}^2 + f_{yy}{}^2 - 2 \cdot f_{xx} \cdot f_{yy} + 4 \cdot f_{xy}{}^2 \tag{32}$$

będącego funkcją monotoniczną w odniesieniu do (28), zatem wprowadzenie progowania dla R jest równoważne progowaniu wyrażenia (32).

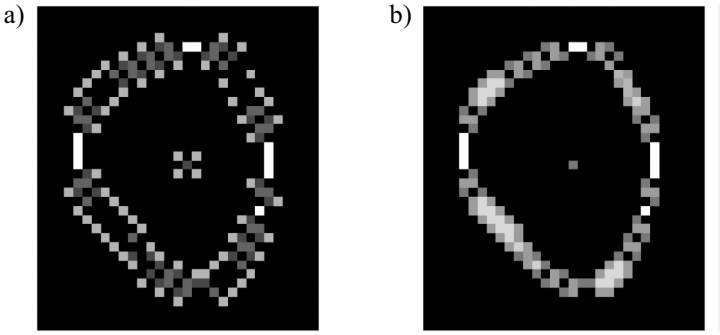
Jak wykazano, wyznaczanie proponowanego parametru (28) w postaci modułu różnicy wartości własnych jest wprawdzie operacją nieliniową, ale jednak stosunkowo prostą w realizacji – porównywalną z przybliżonym wyznaczaniem proporcji wartości własnych opisanym w [7]. Należy podkreślić, iż w przypadku proponowanego parametru R wyliczenie (32) jest dokładne, podczas gdy analogiczne podejście do wyliczenia parametru r daje jedynie wynik przybliżony, który może być uznany za poprawny jedynie dla pewnego zakresu wartości r.

6. Przykładowe wyniki

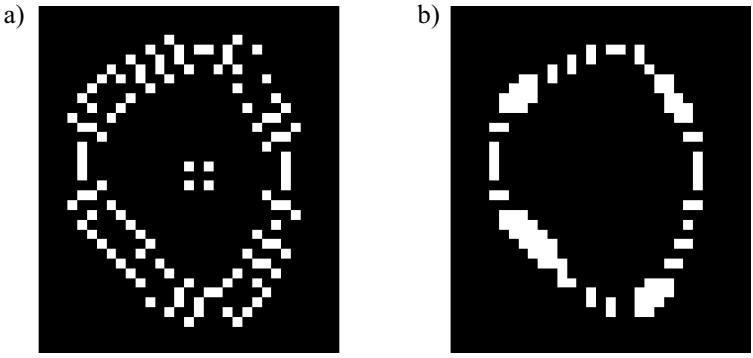
W celu wykazania zalet proponowanego parametru R (28) w porównaniu do współczynnika r (19) opracowano nieduży obraz testowy (rys. 2a) zawierający linię krzywą ułożoną pod różnymi kątami oraz odizolowany punkt ekstremum lokalnego. Współczynniki (19) oraz (28) mają na celu wykazanie obecności grzbietu, przy czym należy pamiętać, iż grzbiet wykrywany za pomocą hesjanu należy traktować jak dwuwymiarowy odpowiednik wklęsłości lub wypukłości funkcji, gdzie miarą tej cechy jest druga pochodna. Zatem w przypadku linii na rysunku 2a grzbiet występuje zarówno na środku widocznej linii, jak i po obu jej stronach.



Rys. 2. Ilustracja opisanej metody wyznaczania grzbietu: a) sztuczny obraz 36×31 pikseli, zawierający linię (czyli trzy równolegle biegnące linie grzbietu) oraz odizolowany punkt ekstremum lokalnego; b) monochromatyczna mapa proponowanych współczynników R (28) (kolor czarny oznacza wartość 0, biały: 344)



Rys. 3. Rozmieszczenie współczynników r (19) dla obrazu testowego z rysunku 2a i dwóch metod wyznaczania hesjanu: a) uśredniającej (13); b) z macierzą pseudoodwrotną (17). Kolor czarny oznacza wartość zero, natomiast biały: a) 13, b) 4



Rys. 4. Wynik binaryzacji dla odpowiednich map współczynników r z rysunku 3. Próg binaryzacji przyjęto na podstawie wartości r dla odizolowanego ekstremum, czyli $r = 1$

Rysunki 3 i 4 eksponują wrażliwość parametru r na metodę wyznaczania hesjanu. Rysunek 2b potwierdza skuteczność proponowanego parametru – parametr R jest identyczny dla obu metod wyznaczania hesjanu, a ponadto umożliwia lokalizację grzbietu zgodnie z oczekiwaniami, podczas gdy parametr r dla każdej z metod wyznaczania hesjanu wykazuje obecność grzbietów jedynie pewnego typu. Przykłady ukazują także, iż wszystkie badane metody wykazują wrażliwość na kierunek badanego grzbietu.

7. Podsumowanie

Jedną z często przyjmowanych interpretacji obecności błędu SE (14) (15) jest założenie, iż treść obrazu jest zaszumiona, a szum ma charakter addytywny. W takim przypadku dopuszczenie modyfikacji wartości centralnej e wydaje się bardziej konsekwentne, niż jej zachowanie bez zmian. Jednak ze względu na popularność masek przedstawionych w równaniu (20) (lub masek do nich zbliżonych [2, 4, 9]), metoda uśredniająca jest często przyjmowana jako właściwe rozwiązanie problemu wyznaczania współczynników szeregu (2). Jak pokazano, rozwiązanie to może być poddane krytyce.

Jedną z istotnych, niekiedy niekorzystnych, cech opisanego podejścia jest wyróżnianie wybranych kierunków – szczególnie pionu i poziomu. W razie potrzeby, można zwiększyć liczbę wyróżnionych kierunków stosując dodatkowe filtry splotowe otrzymane przez obrót macierzy filtrów (20) lub (21) wokół punktu centralnego.

Literatura

[1] Bierski F., *Struktury algebraiczne, Elementy algebry liniowej*. Wyd. 4 poprawione, Kraków, UWND AGH 1999.
[2] Bovik A.(ed.), *Handbook of Image and Video Processing*. Academic Press 2000.
[3] Brown M., Lowe D., *Invariant Features from Interest Points Groups*. Proceedings of the British Machine Vision Conference, Cardiff, Wales, 2002, 656–665.
[4] Hoggar S.G., *Mathematics of Digital Images*. Cambridge University Press 2006.
[5] Kiełbasiński A., Schwetlick H., *Numeryczna algebra liniowa*. Warszawa, WNT 1992.
[6] Lowe D.G., *Object recognition from local scale-invariant keypoints*. International Conference on Computer Vision, Corfu, Greece, 1999, 1150–1157.
[7] Lowe D.G., *Distinctive image features from scale-invariant keypoints*. International Journal of Computer Vision, 60, 2, 2004, 91–110.
[8] Pawlik P., Mikrut S., *Porównanie dokładności metod wyznaczania punktów charakterystycznych na parach zdjęć lotniczych*. Automatyka (półrocznik AGH), t. 11, z. 3, 2007, 141–147.
[9] Tadeusiewicz R., Korohoda P., *Komputerowa analiza i przetwarzanie obrazów*. Kraków, Wydawnictwo Fundacji Postępu Telekomunikacji 1997.
[10] Turowicz A., *Teoria macierzy*. Spisał: W. Mitkowski, Skrypty Uczelniane AGH, Kraków 1985.
[11] Zhao H., Aggarwal J.K., Mandal Ch., Vemuri B.C., *3-D Shape Reconstruction from Multiple Views*. [w:] Bovik A. (ed.), Handbook of Image and Video Processing, Academic Press 2000, 243–257.