

Specyfikacja Algorytmiczna

TECHNIKI REALIZACJI CYFROWEGO PRZETWARZANIA SYGNAŁÓW

Jacek Ambroży Dobrowolski

18 Grudzień 2020

Streszczenie

Celem projektu jest implementacja kilku metod znajdowania punktów do śledzenia oraz algorytmu Optical Flow metodą Lucas-Kanade przy użyciu biblioteki NumPy. I zbadanie wpływu poszczególnych parametrów algorytmu na jego wynik.

Funkcjonalność:

- Wykrywanie dobrych punktów do śledzenia w obrazie
- Odnajdowanie przesunięcia śledzonych narożników

Dane wejściowe:

- Nagranie wideo
- Parametry wybierane z poziomu aplikacji

Dane wyjściowe:

- Strumień wideo z naniesionymi punktami i ich wektorami ruchu
- Podgląd działania poszczególnych etapów przetwarzania

Narzędzia:

- Python
- NumPy
- OpenCV

Wykrywanie dobrych punktów do śledzenia w obrazie

Do wykrywania narożników w obrazie można wykorzystać różne metody. Najbardziej popularne są Harrisa-Stephens[Harris et al., 1988] i Shi-Tomasi[Shi and Tomasi, 1994], są one do siebie podobne i część algorytmu jest wspólna. Obie metody zakładają że narożnik to punkt obrazu w którym intensywność zmienia się w więcej niż dwóch kierunkach. Zmiana w jednym kierunku oznacza krawędź. Niech $I(x, y)$ równa się intensywności piksla o koordynatach x i y . $I(x + u, y + v)$ równa się intensywności piksla odległego o u i v od poprzedniego. Narożnik znajdzie się więc w punkcie gdzie suma kwadratów różnic między tymi wartościami będzie największa. Co można zapisać jako

$$S(x, y) = \sum_u \sum_v w(u, v) [I(x + u, y + v) - I(x, y)]^2$$

gdzie $w(u, v)$ to funkcja okna przesuwanego po obrazie, zwracająca wartość jeden dla małego obszaru i zero dla reszty. Obszarem do zbadania jest działanie innych okien okrągłych, niebinarnych. To równanie możemy przybliżyć wielomianem korzystając z twierdzenia Taylora. Niech I_x i I_y będą pochodnymi I tak aby $I(x + u, y + v) \approx I(u, v) + I_x(u, v)x + I_y(u, v)y$ Daje to wielomian

$$S(x, y) \approx \sum_u \sum_v w(u, v) (I_x(u, v)x + I_y(u, v)y)^2$$

Zapisane w formie macierzowej:

$$S(x, y) \approx \begin{pmatrix} x & y \end{pmatrix} \mathbf{A} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$A = \sum_u \sum_v w(u, v) \begin{bmatrix} I_x(u, v)^2 & I_x(u, v)I_y(u, v) \\ I_x(u, v)I_y(u, v) & I_y(u, v)^2 \end{bmatrix} = \begin{bmatrix} \langle I_x^2 \rangle & \langle I_x I_y \rangle \\ \langle I_x I_y \rangle & \langle I_y^2 \rangle \end{bmatrix}$$

Poprzez analizę wartości własnych macierzy \mathbf{A} można znaleźć narożniki.

- Obie wartości są bliskie zeru dany pixel nie ma dobrych punktów do śledzenia
- Jedna z wartości własnych jest bliska zeru a druga ma dużą wartość dany pixel znajduje się na krawędzi
- Obie wartości własne mają duże dodatnie wartości dany pixel jest narożnikiem

W tym punkcie metody zaczynają się różnić.

Harris-Stephens

Jako że wyznaczanie wartości własnych macierzy jest złożone obliczeniowo. Dlatego stosuje funkcje M_c

$$M_c = \lambda_1 \lambda_2 - \kappa (\lambda_1 + \lambda_2)^2 = \det(A) - \kappa \text{trace}^2(A)$$

Gdzie κ to stała wyznaczana empirycznie, która odpowiada sa czułość algorytmu. W literaturze można się spotkać z wartościami między 0.04 do 0.15.

Shi-Tomasi

Metoda Shi-Tomasi bezpośrednio wyznacza $\min(\lambda_1, \lambda_2)$ Co daje nieco lepsze wyniki kosztem złożoności obliczeniowej.

Metoda FAST

Metoda FAST (Features from accelerated segment test)[Rosten and Drummond, 2006] polega na znalezieniu punktów dla których w danym promieniu. Znaczna część pixeli różni się od centralnego pixela o wartość progu detekcji. Zaletą tej metody jest niska złożoność obliczeniowa. Algorytm ten ma 3 parametry które można zbadać.

Estymacja przesunięcia śledzonych narożników

Metoda Lucas-Kanade[Lucas and Kanade, 1981] jest szeroko stosowana w wizji komputerowej. Polega na założeniu że ruch w sąsiednich pixeli przebiega w jednym kierunku. Jako metoda o działaniu lokalnym nie radzi sobie z wyznaczaniem ruchu dla regionów o jednakowej intensywności. Dlatego jako dane wejściowe podajemy dobre punkty do śledzenia. Metoda ta zakłada również nie wielki ruch między poszczególnymi klatkami. Wynikają z tego pewne ograniczenia takie jak słabe działanie dla szybkich obiektów lub konieczność szybszego próbkowania obrazu. Algorytm wyznacza wektor który najlepiej opisuje ruch danego wycinka obrazu.

$$\begin{aligned} I_x(q_1)V_x + I_y(q_1)V_y &= -I_t(q_1) \\ I_x(q_2)V_x + I_y(q_2)V_y &= -I_t(q_2) \\ &\vdots \\ I_x(q_n)V_x + I_y(q_n)V_y &= -I_t(q_n) \end{aligned}$$

Gdzie V_x to współrzędna pozioma wektora ruchu regionu a V_y to współrzędna pionowa tegoż wektora. q_1, q_2, \dots, q_n to pixele wewnątrz regionu. A $I_x(q_i), I_y(q_i), I_t(q_i)$ to pochodne cząstkowe intensywności obrazu. Równanie to można zapisać w formie macierzowej jako. $\mathbf{A}v = b$ gdzie:

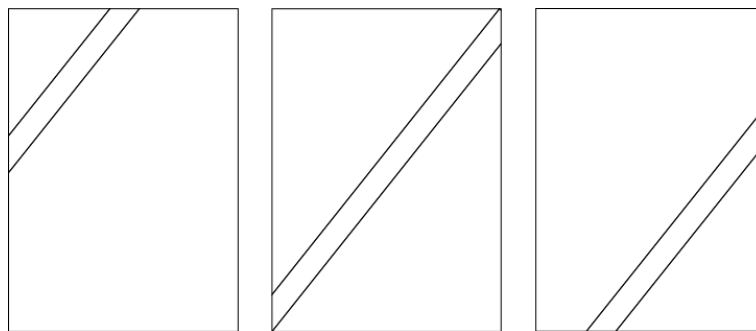
$$\mathbf{A} = \begin{bmatrix} I_x(q_1) & I_y(q_1) \\ I_x(q_2) & I_y(q_2) \\ \vdots & \vdots \\ I_x(q_n) & I_y(q_n) \end{bmatrix} \quad v = \begin{bmatrix} V_x \\ V_y \end{bmatrix} \quad b = \begin{bmatrix} -I_t(q_1) \\ -I_t(q_2) \\ \vdots \\ -I_t(q_n) \end{bmatrix}$$

Jest to zwykle układ nadokreślony metoda Lucas-Kanade przybliża wynik metodą najmniejszych kwadratów.

$$v = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T b$$

Problem przesłony

Metoda Lucas Kanade jest podatna na problem przesłony jest to zjawisko kiedy określenie kierunku ruchu nie jest możliwe na podstawie jego drobnego wycinka. Efekt jest zobrazowany na poniższej ilustracji.



Ilustracja obrazująca efekt przysłony

Piramidy i iteracje

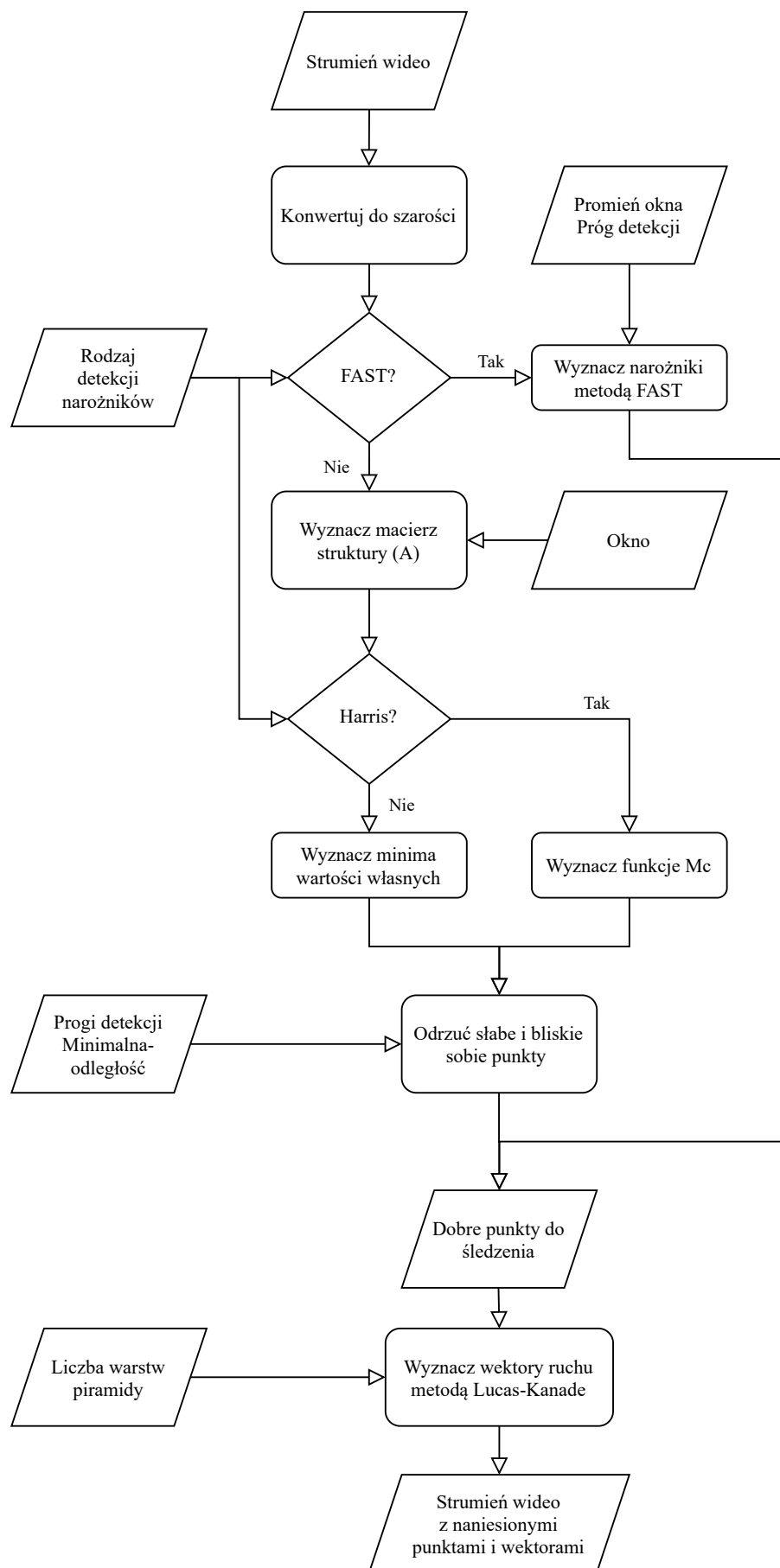
W literaturze można spotkać się z określeniem piramid. Jest to porostu metoda poprawienia działania algorytmu Lucas-Kanade poprzez obliczanie wektora najpierw dla obniżonej rozdzielczości. A następnie dla coraz większej. Ma to na celu zmniejszenie szumu i losowości w miejscach gdzie wyznaczenie przesunięcia nie jest możliwe. Można spodziewać się poprawienia trafności algorytmu stosując piramidy.

Badanie działania algorytmów

Najprostszą miarą badającą złożoność obliczeniową są uzyskane klatki na sekundę. Wadą takiego pomiaru jest duży wpływ implementacji. Inne miary jakich planuje użyć są empiryczne. Istnieją specjalne zestawy danych rzeczywistych i syntetycznych, jednak porównanie działania algorytmów na nich zostało już wykonane.

Dalszy rozwój projektu

Projekt może zostać rozbudowany o prostą stabilizację obrazu. Opartą na wygładzaniu ruchu kamery poprzez odkształcanie obrazu.



Literatura

- [Harris et al., 1988] Harris, C. G., Stephens, M., et al. (1988). A combined corner and edge detector. In *Alvey vision conference*, volume 15, pages 10–5244. Citeseer.
- [Lucas and Kanade, 1981] Lucas, B. and Kanade, T. (1981). An iterative image registration technique with an application to stereo vision (ijcai). volume 81.
- [Rosten and Drummond, 2006] Rosten, E. and Drummond, T. (2006). Machine learning for high-speed corner detection. In Leonardis, A., Bischof, H., and Pinz, A., editors, *Computer Vision – ECCV 2006*, pages 430–443, Berlin, Heidelberg. Springer Berlin Heidelberg.
- [Shi and Tomasi, 1994] Shi, J. and Tomasi, C. (1994). Good features to track. pages 593–600.