

Marek SZCZEPAŃSKI
Politechnika Śląska

FILTRY CZASOWE I PRZESTRZENNE W PRZETWARZANIU STRUMIENI WIDEO

Streszczenie. Artykuł ten dokonuje przeglądu istniejących technik filtracji sekwencji wideo, ze szczególnym uwzględnieniem możliwości ich zastosowania do aplikacji czasu rzeczywistego. Zaproponowano też rozwiązanie wykorzystujące detekcję ruchu w kolejnych klatkach ograniczającą efekt smużenia, połączone z zastosowaniem przestrzennych wersji wektorowej mediany oraz filtru FDPA. Szczególnie efektywne jest zastosowanie proponowanych rozwiązań w połączeniu z poprawą kontrastu do sekwencji wykonywanych przy słabym oświetleniu.

SPATIAL AND TEMPORAL FILTERS IN VIDEO PROCESSING

Summary. An overview of real-time video filtering techniques was presented in this paper. The new approach is based on temporal mean combined with movement detection and spatial filtering was presented. Presented technique provides excellent noise suppression ability with low computational complexity.

1. Wprowadzenie

Ciągły rozwój procesów technologicznych związanych z produkcją sensorów obrazów jest determinowany dużą podażą na rynku urządzeń multimedialnych. Niski koszt produkcji matrycy sprawia, że są one montowane nie tylko w aparatach cyfrowych i kamerach, ale również komputerach przenośnych czy kamerach internetowych, umożliwiając prowadzenie rozmów wideo. Producenci sprzętu przenośnego, takiego jak telefony komórkowe, coraz częściej wyposażają swoje produkty w urządzenia do przechwytywania obrazu. Rozwój technologii UMTS wymusza często zastosowanie dwóch przetworników w jednym urządzeniu, tak aby możliwe było robienie zdjęć i prowadzenie rozmów wideo przy użyciu jednego wyświetlacza.

Niestety, wysoki poziom miniaturyzacji sensorów pociąga za sobą bardzo niską jakość rejestrowanego materiału. Ponadto zwiększanie liczby megapikseli, upakowanych na tak małej powierzchni w celu poprawienia jakości obrazu, powoduje czasem efekt przeciwny do zamierzonego. Wynika to z konieczności naświetlenia matrycy, podobnie jak błony światłoczułej. Im mniejsza powierzchnia jednego piksela, tym bardziej sygnał przez niego generowany należy wzmocnić. Duża część małych

i tanich przetworników daje bardzo ciemny obraz w warunkach słabego oświetlenia. Wzmacniając sygnał, potęgujemy również zakłócenia, którymi jest on obciążony.

Niezbędne jest więc zastosowanie algorytmów poprawy jakości tak rejestrowanego obrazu. Szybkość wykonywania tych procesów jest tutaj parametrem krytycznym. Czas ich realizacji powinien być mniejszy niż odstęp pomiędzy kolejnymi klatkami.

Celem pracy jest dokonanie przeglądu metod filtracji czasowej i przestrzennej pod kątem ich przydatności do obróbki strumieni wideo w czasie rzeczywistym. W pracy omówione zostały trzy grupy filtrów: filtry przestrzenne, czasowe oraz czasowo-przestrzenne. W ramach tej ostatniej grupy zaproponowano rozwiązania umożliwiające przetwarzanie w czasie rzeczywistym, zachowujące równocześnie wysoką efektywność filtracji.

2. Filtry przestrzenne

Filtracja obrazów ruchomych bazuje na pozyskiwaniu i wykorzystywaniu informacji zawartych we wszystkich klatkach filmu. Niestety, większość czasowych algorytmów usuwających szumy powoduje rozmazywanie obrazów lub powstawanie nieciągłości w postaci poszarpanych krawędzi. Dzięki zastosowaniu filtracji przestrzennej w postprocessingu możemy wyeliminować wspomniane zniekształcenia. Natomiast uzyskanie informacji o krawędziach czy operacje na histogramie w preprocesingu są niezbędne do efektywnego działania algorytmów w funkcji czasu.

W ramach testów zaimplementowano następujące liniowe filtry przestrzenne:

- przetwarzanie splotowe,
- mediana skalarna,
- mediana wektorowa (VMF) [1].

3. Filtry czasowe

Szumy wprowadzane przez matryce CCD i CMOS w znacznym stopniu obniżają jakość rejestrowanego materiału i powodują duże straty podczas kompresji. Dzięki temu, że mamy do czynienia z sekwencją obrazów, a nie z pojedynczą klatką, możemy wyeliminować zakłócenia. Istotny jest fakt, że szum przetwornika cechuje się niską korelacją pomiędzy poszczególnymi klatkami, podczas gdy elementy obrazu, nawet szybko zmieniającego się, są skorelowane.

Właściwości te wykorzystywane są podczas filtracji czasowej, gdy stosowane są różne warianty uśredniania wartości poszczególnych pikseli w kolejnych klatkach sekwencji wideo.

Najprostszą realizacją filtracji czasowej jest *czasowy filtr uśredniający* (ang. *Temporal Arithmetic Mean Filter - TAMF*); wynik działania takiego filtra można przedstawić następującymi zależnościami:

$$\hat{F}(i, j, t) = \frac{1}{n} \sum_{\Delta t=0}^{n-1} F(i, j, t + \Delta t) \quad (1)$$

lub

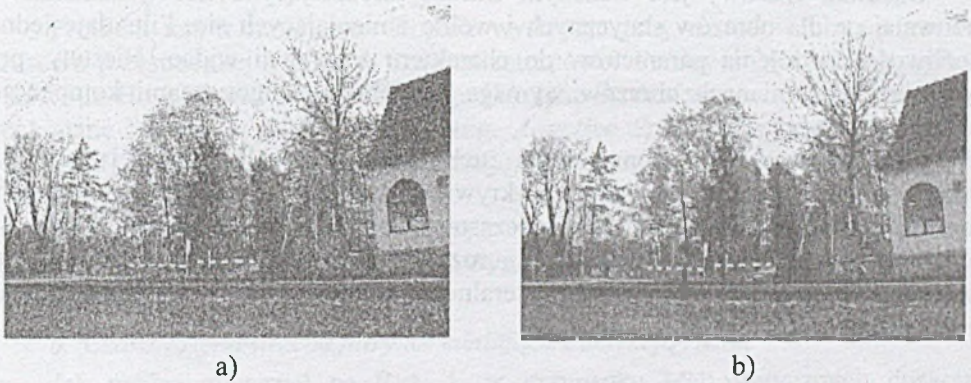
$$\hat{F}(i, j, t) = \frac{1}{2n+1} \sum_{\Delta t=-n}^n F(i, j, t + \Delta t) \quad (2)$$

gdzie: F, \hat{F} oznaczają odpowiednio klatkę wejściową i wynik filtracji, a n określa rozmiar okna czasowego. Przy czym zależność (2) opisuje filtr symetryczny względem przetwarzanej ramki.

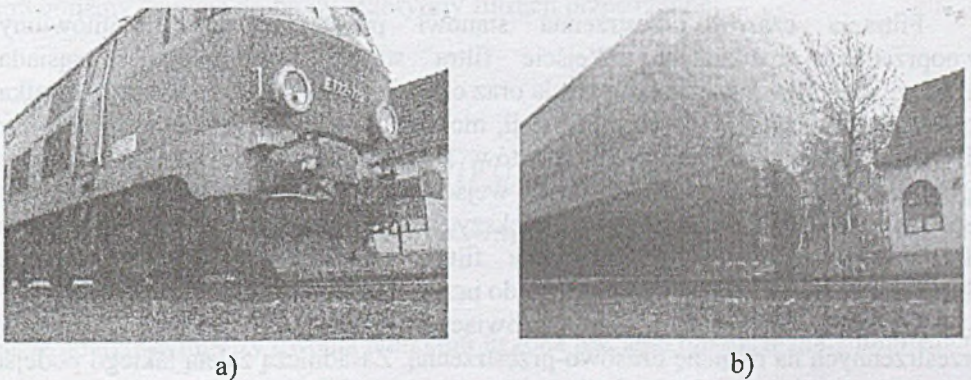
Sposób ten, choć najprostszy i najszybszy, nadaje się tylko do sekwencji statycznych, ponieważ uśrednienie klatek, na których znajdują się obiekty w ruchu, powoduje powstawanie na przetworzonym obrazie efektu smużenia (ang. *ghosting*).

Sceny niezmienną się w czasie są znacząco poprawiane. Częściowo eliminowany jest wpływ pojedynczych pikseli, których wartość odbiega od wartości ich odpowiedników w innych obrazach, branych pod uwagę podczas uśredniania. Jednak jeżeli dany piksel na jednym obrazie znacząco różni się od swoich odpowiedników w pozostałych obrazach, jego wpływ będzie widoczny, ponieważ jego wartość będzie miała duże znaczenie podczas sumowania w równaniach (1) i (2).

Przykłady skutecznej filtracji mocno zaszumianej sekwencji dla sceny statycznej przedstawiono na rysunku 1, kolejna ilustracja (Rys. 2) przedstawia zaś działanie filtra dla sceny dynamicznej.



Rys. 1. Obraz statyczny, 30 klatek na sekundę:
a) klatka n , b) wynik uśrednienia klatek $(n-9) \dots n$



Rys. 2. Obraz dynamiczny, 30 klatek na sekundę:
a) klatka n , b) wynik uśrednienia klatek $(n-9) \dots n$

Problem rozmywania obiektów ruchomych można rozwiązać przez wprowadzenie progów minimalnej zmiany wartości między klatkami i ograniczyć przetwarzanie czasowe do tych właśnie obszarów, tak więc wyjście filtra będzie miało postać:

$$\hat{F}(i, j, t) = \begin{cases} \frac{1}{n} \sum_{\Delta t=0}^{n-1} F(i, j, t + \Delta t), & \text{dla } \sum_{\Delta t=0}^{n-1} |F(i, j, t + \Delta t) - F(i, j, t + \Delta t - 1)| \leq \beta \\ F(i, j, t), & \text{w przeciwnym przypadku.} \end{cases} \quad (3)$$

Zależność (3) może zostać zmodyfikowana w taki sposób, by w przypadku scen elementów ruchomych zastosować filtrację przestrzenną.

Analogicznie do filtrów średniej czasowej można zdefiniować czasowy filtr Gaussa. Do realizacji czasowej wersji Gaussianu wykorzystywana jest jedno-wymiarowa funkcja Gaussa:

$$\hat{F}(i, j, t) = \sum_{t=-n}^{t=n} g(\sigma, t) * F(i, j, t) \quad (4)$$

Gaussian czasowy jest ważonym filtrem uśredniającym. Efekty działania są zadowalające dla obrazów statycznych i wolno zmieniających się. Filtr daje jednak możliwość dostrojenia parametrów do charakteru sekwencji wideo. Niestety, przy dynamicznych zmianach obrazów wymaga uzupełnienia algorytmami kompensacji ruchu.

Filtrację czasową z kompensacją ruchu zaproponowali Dubois i Sabri [3]. Ogólna idea kompensacji polega na wykrywaniu przemieszczenia każdego elementu obrazu i zastosowaniu jednego z filtrów czasowych wzdłuż trajektorii ruchu.

Innym rozwiązaniem problemu rozmywania ruchomych obiektów jest zastosowanie czasowej wersji filtra bilateralnego, którego wersja przestrzenna opisana jest w pracach [9] oraz [4].

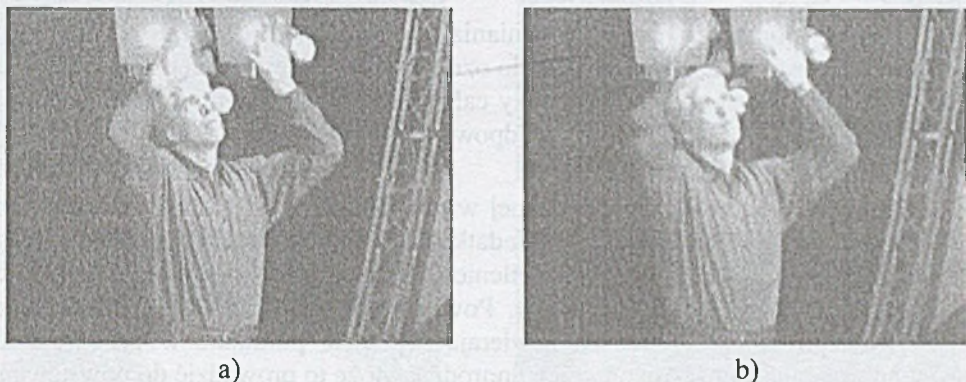
4. Filtry czasowo-przestrzenne

a. Przegląd rozwiązań

Filtracja czasowo-przestrzenna stanowi połączenie metod omówionych w poprzednich rozdziałach. Wejście filtra stanowi informacja o sąsiadach przestrzennych przetwarzanego piksela oraz o jego odpowiednikach w innych klatkach sekwencji. Korzystając z tych informacji, możemy skutecznie poprawić jakość filmu, nie wprowadzając negatywnych efektów filtracji albo w znacznym stopniu je ograniczyć. Większy zbiór informacji wejściowych potrzebnych do przetworzenia jednego piksela oznacza również większy nakład obliczeniowy, między innymi dlatego omawiane w tym rozdziale filtry należą do najbardziej złożonych obliczeniowo i najczęściej stosuje się je do przetwarzania 'off-line'.

Powstała cała grupa filtrów stanowiących uogólnienie standardowych filtrów przestrzennych na domenę czasowo-przestrzenną. Zasadniczą zaletą takiego podejścia jest brak konieczności detekcji czy też kompensacji ruchu, przy znacznym ograniczeniu rozmywania obiektów w ruchu. Do tych metod zaliczyć możemy

trójwymiarowe filtry medianowe opisane między innymi w pracach [5, 10, 11]. Inne rozwiązania stosowane do barwnych sekwencji wykorzystują trójwymiarową wersję mediany wektorowej (*VMF3D*). Metody te, niestety, powodują często powstawanie poszarpanych krawędzi ruchomych obiektów (Rys. 3).



Rys. 3. Klatka z sekwencja testowa 'Żongler' oraz wynik działania filtru *VMF3D* (okno $3 \times 3 \times 3$, norma L_1)

Problem ten w mniejszym stopniu występuje przy zastosowaniu rozszerzonej czasowo-przestrzennej wersji dyfuzji anizotropowej Perony i Malika [6], opisanej przez Suk Ho Lee oraz Moon Gi Kanga [7]. Inne bardzo skuteczne rozwiązanie wykorzystane zostało w filtrze *ASTA* (ang. *Adaptive Spatio-Temporal Accumulation Filter*) [2], wykorzystującym kombinację czasowego oraz przestrzennego filtru bilateralnego [9].

Filtr *ASTA* oraz czasowo-przestrzenna dyfuzja anizotropowa są bardzo skuteczne w redukcji szumów, wymagają jednak znacznych nakładów obliczeniowych.

b. Czasowo-przestrzenne filtry uśredniające z detekcją ruchu

Jak można zobaczyć na Rys. 1, w przypadku scen statycznych doskonale rezultaty uzyskamy, stosując proste uśrednianie w czasie, które jednak kompletnie nie sprawdza się, gdy w obrazie pojawiają się obiekty ruchome (Rys. 2); wówczas powinniśmy zaprząć do pracy algorytmy filtracji przestrzennej.

$$\hat{F}(i, j, t) = \begin{cases} \frac{1}{n} \sum_{\Delta t=0}^{n-1} F(i, j, t + \Delta t), & \text{dla} \\ \sum_{\Delta t=0}^{n-1} \text{dist}(F(i, j, t + \Delta t), F(i, j, t + \Delta t - 1)) \leq \beta & \\ F_{\text{filt}2D}(i, j, t), & \text{w przeciwnym przypadku,} \end{cases} \quad (5)$$

gdzie odległość między dwoma punktami liczona jest jako moduł różnicy luminancji:

$$\text{dist}(F_1, F_2) = \left| \frac{1}{3} [(F_{1R} + F_{1G} + F_{1B}) - (F_{2R} + F_{2G} + F_{2B})] \right| \quad (6)$$

W zależności (5) dla elementów ruchomych wyznaczany jest wynik filtracji przestrzennej. W naszym przypadku przetestowaliśmy dwa rozwiązania: wektorową medianę (*VMF*), dalej oznaczaną jako *TAMF-SVMF*, oraz filtr oparty na idei ścieżek cyfrowych *FDPA* (ang. *Fast Digital Paths Approach*) [230], oznaczany jako *TAMF-SFDPA*.

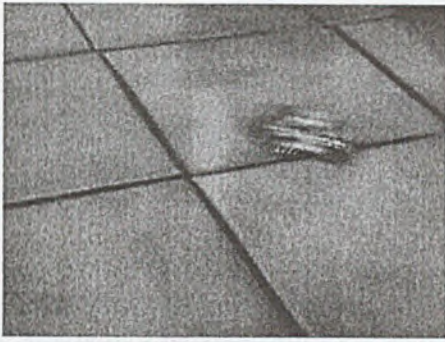
Wynik działania czasowego uśredniania z pominięciem elementów ruchomych przedstawiony jest na rysunku 4b: na biało oznaczono piksele przeznaczone do filtracji przestrzennej. Zaznaczone piksele zostały całkowicie pominięte przez filtr, ponieważ różnice między luminancjami ich odpowiedników w kolejnych klatkach były zbyt duże.

W sekwencji testowej wykorzystanej w przykładzie z rysunku 4 przedstawiono bardzo szybko poruszający się obiekt. Dodatkowo w trakcie przetwarzania wstępnego z uwagi na bardzo słabe warunki oświetleniowe konieczne było rozjaśnienie obrazu, co spowodowało wzmocnienie szumów. Powyższy przykład pokazuje pewną wadę przedstawionego podejścia: maska zawierająca piksele pomijane w przetwarzaniu czasowym często jest poszarpana i niejednorodna. Może to prowadzić do powstawania zauważalnych artefaktów w otoczeniu ruchomych obiektów. Rozwiązaniem tego problemu może być przeprowadzenie operacji zamknięcia na naszej masce. Wyniki takiej operacji przedstawione są na Rys. 4 c oraz 4 d.

Średnia szybkość działania poszczególnych algorytmów dla sekwencji testowych 'Autko' oraz 'Żongler' o rozdzielczości 320x240 przedstawiona jest w tabeli 1. Testy zostały przeprowadzone na komputerze z procesorem *Intel Core 2 Quad Q6600* pracującym z częstotliwością 2,4 GHz, przy wykorzystaniu wszystkich czterech rdzeni procesora. Wszystkie filtry bazujące na czasowym uśrednianiu korzystają z okna o długości 5 klatek, w filtrach z detekcją ruchu przyjęto próg $\beta = 5$. Szybkość działania filtrów uśredniających porównano z czasowo-przestrzenną wersją wektorowej mediany (*VMF3D*) dla okna filtracji o rozmiarze 3x3x3 oraz filtrami przestrzennymi: wektorową medianą (*VMF*) i filtrem opartym na ścieżkach cyfrowych o długości 2 (*FDPA*).

Tabela 1
Średnia szybkość działania wybranych filtrów dla sekwencji 'Autko' oraz 'Żongler'

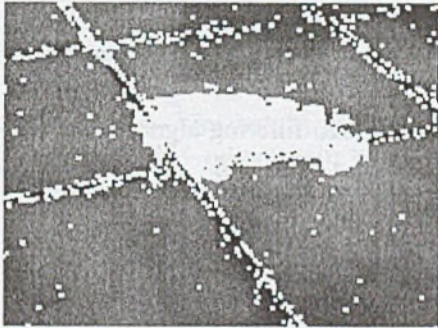
	'Autko' [fps]	'Żongler' [fps]
<i>TAMF-SVMF</i>	109	80,5
<i>TAMF-SVMF</i> +closing	71	57,3
<i>TAMF-FDPA</i>	85,1	47,7
<i>TAMF</i>	168,7	148,7
<i>VMF3D 3x3x3</i>	9,7	9,5
<i>VMF 3x3 L1</i>	57,3	56,1
<i>FDPA1st 2steps</i>	11,8	11,7



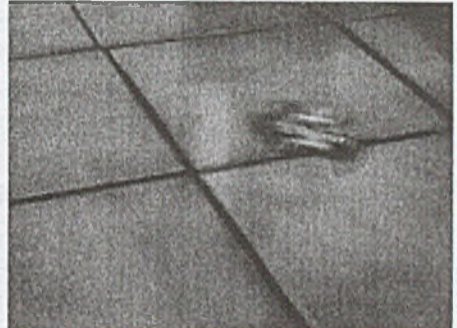
a)



b)



c)



d)

Rys. 4. Przykład działania połączenia czasowego filtru uśredniającego dla $\beta = 5$, $n = 5$ i wektorowej mediany (okno 3×3 norma L_1): a) oryginalna klatka z sekwencji 'Autko', b) filtracja czasowa – maska pikseli pomijanych, c) maska pikseli pomijanych po operacji zamknięcia z elementem strukturalnym 3×3 , d) końcowy wynik filtracji

5. Podsumowanie

W niniejszej pracy dokonano przeglądu istniejących technik przetwarzania sekwencji wideo, ze szczególnym uwzględnieniem możliwości ich zastosowania do aplikacji czasu rzeczywistego. Zaproponowano rozwiązanie wykorzystujące detekcje ruchu w kolejnych klatkach ograniczające efekt smużenia, połączone z zastosowaniem przestrzennych wersji wektorowej mediany oraz filtru *FDPA*.

Uzyskane wyniki pokazują wysoką wydajność zaproponowanych filtrów - są one wyraźnie szybsze nawet od rozwiązań wykorzystujących filtrację przestrzenną, a przy tym zapewniają bardzo skuteczną redukcję zakłóceń.

Szczególnie efektywne jest zastosowanie proponowanych rozwiązań w połączeniu z poprawą kontrastu do sekwencji wykonywanych przy słabym oświetleniu, np. z systemu monitoringu.

BIBLIOGRAFIA

1. Astola J., Haavisto P., Neuvo Y.: Vector Median Filters. IEEE Proc., volume 78, p. 678 – 689, 1990.

2. Bennett P.E., McMillan L.: Video Enhancement Using Per-Pixel Virtual Exposures. ACM SIGGRAPH 2005.
3. Dubois E., Sabri S.: Noise Reduction in Image Sequences Using Motion-Compensated Temporal Filtering. Communications, IEEE Transactions on vol.32 no.7, p. 826-831, 1984.
4. Durand F., Dorsey J.: Fast Bilateral Filtering for the Display of High-Dynamic-Range Images. ACM SIGGRAPH 2002.
5. Meguro M., Taguchi A., Hamada N.: Data-dependent weighted median filtering with robust motion information for image sequence restoration. Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on, vol.2, p. 424-428, 1999.
6. Perona P., Malik J.: Scale-space and edge detection using anisotropic diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol 12, p. 629-639.
7. Suk Ho Lee, Moon Gi Kang: Spatio-temporal video filtering algorithm based on 3-D anisotropic diffusion equation. Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on, vol. 2, p. 447-450, 1998.
8. Szczepanski M., Smolka B., Plataniotis K. N., Venetsanopoulos A. N.: On the Geodesic Paths Approach to Color Image Filtering. Signal Processing, vol. 83, no. 6, p. 1309-1342, 2003.
9. Tomasi C., Manduchi R.: Bilateral Filtering for Gray and Color Images, ICCV, p. 839 – 846, 1998.
10. Viero T., Neuvo Y.: Non-moving regions preserving median filters for image sequence filtering. Systems Engineering, 1991., IEEE International Conference on, p. 245-248, 1991.
11. Viero T., Oistamo K., Neuvo Y.: Three-dimensional median-related filters for color image sequence filtering. IEEE Transactions On Circuits And Systems For Video Technology, vol. 4, no. 2, p. 129-142, 1994.

Recenzent: Dr hab. inż. Wiesław Kotarski

Abstract

From several years we can observe increasing interest in video processing. Video noise reduction without structure degradation is perhaps the most challenging video enhancements task. Several techniques have been proposed over the years. Among them are standard noise reduction techniques, the so-called spatial filters, applied to subsequent frames of the video stream. However, standard image processing techniques cannot utilize all available information i.e. similarities in neighboring frames, so modern video denoising algorithms utilize also temporal information .

An overview of real-time video filtering techniques was presented in this paper. The new approach is based on temporal mean combined with movement detection and spatial filtering was presented. Presented technique provides excellent noise suppression with low computational complexity.