

Andrzej POPLAWSKI
Uniwersytet Zielonogórski
Instytut Metrologii, Elektroniki i Informatyki

EFEKTYWNOŚĆ SYSTEMU NADZORU WIZYJNEGO W WARUNKACH CZĘSTEJ ZMIANY SCENY

W urządzeniach Internetu rzeczy (IoT) coraz częściej stosowane są systemy monitorowania i nadzoru wizyjnego. Wykorzystuje się tu różne techniki kompresji, które przyczyniają się do pogorszenia jakości obrazów. Jakość ta zależy nie tylko od zastosowanej kamery i techniki kompresji, ale również od treści rejestrowanej przez kamerę. W artykule zostaną przedstawione badania efektywności kompresji systemu nadzoru wizyjnego w warunkach częstej zmiany sceny.

THE EFFECTIVENESS OF THE VIDEO SURVEILLANCE SYSTEMS IN CONDITIONS OF FREQUENT SCENE CHANGE

Monitoring and video surveillance systems are increasingly used in IoT devices. Various compression techniques are used here, which contribute to the deterioration of image quality. This quality depends not only on the camera and compression technique used, but also on the content recorded by the camera. The article presents research on the coding efficiency of video surveillance system in the conditions of frequent scene change.

1. WPROWADZENIE

W urządzeniach Internetu rzeczy (IoT) coraz częściej stosowane są różnego rodzaju systemy monitorowania i nadzoru wizyjnego. W celu efektywnego przesyłania i gromadzenia danych wizyjnych, stosuje się kompresję, która ma na celu zmniejszenie ilości potrzebnych do przesłania bądź gromadzenia danych [1, 2]. Kompresja przyczynia się jednak do pogorszenia jakości obrazów. Jakość ta zależy nie tylko od zastosowanej techniki kompresji, ale również od zarejestrowanej treści.

Techniki kompresji sekwencji wizyjnych są stosowane od wielu lat i stają się coraz efektywniejsze. W ostatnich latach opracowano wiele standardów kodowania, takich jak H.263, H.264/AVC [3, 4], powszechnie stosowanych w wielu dziedzinach. Ostatnie prace zaowocowały nową, wydajną technologią kompresji wideo High Efficiency Video Coding - HEVC [5, 6]. HEVC to bardzo duże osiągnięcie w dziedzinie kompresji wideo, wykorzystywane również w IoT [2]. Zastosowanie techniki HEVC zamiast techniki AVC, pozwala na około dwukrotną redukcję prędkości transmisji [7, 8].

Wykorzystanie podobieństw między sąsiadującymi ramkami ma kluczowe znaczenie w kompresji sekwencji wizyjnych, ich występowanie jest warunkiem koniecznym zapewnienia wysokiej wydajności kompresji. Im większe podobieństwo sąsiednich klatek, tym bardziej wydajna może być kompresja. W praktyce stopień podobieństwa sąsiednich klatek nie jest stały i zmienia się w zależności od zawartości materiału wideo. Oczywiście jest, że od czasu do czasu w sekwencji wizyjnych następuje zmiana sceny. Taka sytuacja może mieć miejsce np. po zmianie kamery rejestrującej. W tej sytuacji podobieństwa między sąsiadującymi ramkami są na ogół niewielkie. W rzeczywistości na prawdziwym materiale zmiana sceny nie występuje często. Jednak im szybszy jest ruch kamery, im więcej dynamicznego ruchu w kadrze kamery, tym bardziej niepodobne stają się sąsiednie klatki. Również szybkozmienne tło, np. falująca powierzchnia wody, latające konfetti, pokazy sztucznych ogni czy wyładowania atmosferyczne, mogą powodować znaczne różnice między sąsiednimi ramkami. W takich przypadkach często mówi się o zmianie sceny. W systemach nadzoru wizyjnego materiał z dużą ilością ruchu zwykle jest cenniejszy niż statyczne nagranie.

W literaturze kwestia zmiany sceny w sekwencji wizyjnych jest traktowana głównie jako problem wykrywania ujęć w filmie. Odbywa się to w celu segmentacji i klasyfikacji materiałów wideo. Podział sekwencji na ujęcia służy m.in. do indeksowania i archiwizacji danych wizyjnych [9–11].

2. METODOLOGIA BADAŃ EKSPERYMENTALNYCH

Głównym celem badań eksperymentalnych było zbadanie zachowania kodeka wizyjnego HEVC w warunkach częstej zmiany sceny. Badania przeprowadzono przy użyciu reprezentatywnej grupy testowych sekwencji wizyjnych: Basket, City, Crew, Harbour, Ice i Soccer formacie CIF 30Hz (352x288) i długości 6,4 sekundy (192 obrazy) dostępnych w [12]. Wybrane sekwencje są powszechnie używane do oceny skuteczności kompresji sekwencji wizyjnych, charakteryzują się zmienną dynamiką zarówno na pierwszym planie, jak i w tle. W badaniach nad kodekami wizyjnymi dostępnymi w literaturze stosuje się sekwencje wizyjne bez zmiany sceny. Sekwencje wizyjne zostały zatem zmodyfikowane w celu uzyskania symulowanej zmiany sceny z określoną częstotliwością. W artykule wprowadzono parametr SC (ang. Scene Change). Jego wartość określa, jak często występuje zmiana sceny w wejściowej sekwencji wizyjnej. Założono, że w badaniach zostaną wykorzystane sekwencje wizyjne, dla których parametr SC otrzyma wartości: 32, 16, 8, 4, 2, 1. W związku z tym najrzadziej zmiana sceny będzie następować co 32 obrazy, następnie co 16 obrazów, a najczęściej co 1 obraz. Zatem dla wartości parametru SC równej 1, każda klatka sekwencji wizyjnej będzie różna od następnej lub poprzedniej klatki w tej sekwencji. Ponadto założono, że dla wartości parametru SC równej 0 zmiana sceny nie występuje. Kodowanie wideo w tym przypadku będzie wykonywane przy użyciu testowych sekwencji wizyjnych bez zmiany sceny.

Aby zasymulować zmianę sceny, oryginalna testowa sekwencja wizyjna składająca się ze 192 obrazów została podzielona na dwa równe fragmenty. Pierwsza część to kolejne obrazy od 1. do 96., druga część to kolejne obrazy od 97. do 192. W następnym kroku obrazy z pierwszej i drugiej części zostały połączone na przemian. Liczba kolejnych połączonych obrazów każdej części zależy od wartości parametru SC. Na przykład, dla zmiany sceny co 4 obrazy (SC=4), w wynikowej sekwencji zapisywane były kolejno obrazy od 1. do 4., następnie obrazy od 97. do 100., następnie od 5. do 8., dalej od 101. do 104. i tak dalej. Kolejne obrazy pochodzą z oryginalnej sekwencji wizyjnej. Przy takiej metodzie co cztery obrazy następuje symulowana zmiana sceny. Obrazy 4. i 5., a następnie 8. i 9. nowej sekwencji są do siebie niepodobne. Możemy zatem powiedzieć, że nastąpiła zmiana sceny. Taki sposób zasymulowania zmiany sceny powoduje, że charakter nowo utworzonej sekwencji (dynamika, kolory) zostaje zachowany. Dla każdej oryginalnej sekwencji testowej utworzono sześć zmodyfikowanych sekwencji testowych wizyjnych dla wartości parametru SC od 32 do 1.

W badaniach wykorzystano powszechnie dostępne oprogramowanie referencyjne HEVC [13]. Zastosowano konfigurację kodowania dostępu losowego, zgodnie z definicją w [14]. Jako miarę jakości zastosowano współczynnik PSNR. Eksperyment polegał na kodowaniu, a następnie dekodowaniu sekwencji wizyjnych z dwunastoma różnymi prędkościami transmisji: 128, 196, 256, 384, 512, 640, 768, 896, 1024, 1152, 1280, 1408 kb/s oraz wyznaczeniu wartości PSNR. W artykule wprowadzono miarę różnicy wartości PSNR – VSPD, zdefiniowaną jako:

$$VSPD = PSNR_0 - PSNR_{SC} [dB], \quad (1)$$

gdzie:

$PSNR_0$ - wartość współczynnika PSNR w przypadku braku zmiany sceny,

$PSNR_{SC}$ - wartość współczynnika PSNR w przypadku zmiany sceny.

Ta miara określa różnicę między wartością PSNR w przypadku braku zmiany sceny, a wartością PSNR w przypadku, gdy zmiana sceny występuje.

3. WYNIKI BADAŃ EKSPERYMENTALNYCH

W tabeli 1 przedstawiono uzyskane wartości PSNR dla testowanych sekwencji wizyjnych oraz wybranych, niektórych prędkości transmisji. Zaobserwowany największy spadek wartości PSNR odnotowano dla sekwencji City - średnio 2,13 dB, natomiast najmniejszy spadek wartości PSNR zaobserwowano w przypadku sekwencji Crew - średnio 0,25 dB.

Tabela 1

Wybrane wartości współczynnika PSNR dla badanych wartości parametru SC

Nazwa sekwencji	Prędkość transmisji [kb/s]	Wartość PSNR [dB] dla danej wartości parametru SC						
		0	32	16	8	4	2	1
Basket	256	25,98	25,61	24,95	24,89	25,02	24,67	23,06
	512	29,43	29,15	28,50	28,50	28,56	28,21	25,60
	768	31,75	31,50	30,84	30,90	30,97	30,63	27,50
City	256	35,41	34,96	33,23	33,46	33,71	33,33	29,82
	512	38,70	38,36	36,94	37,10	37,27	37,07	32,45
	768	40,67	40,42	39,17	39,36	39,38	39,14	34,14
Crew	256	34,42	34,57	34,43	34,30	34,47	34,09	32,41
	512	37,22	37,52	37,34	37,30	37,38	37,12	35,19
	768	38,98	39,32	39,17	39,12	39,16	38,88	37,06
Harbour	256	29,23	28,92	28,29	28,42	28,47	28,34	26,02
	512	31,90	31,64	31,17	31,18	31,21	31,10	28,49
	768	33,65	33,44	32,99	32,97	32,95	32,83	30,22
Ice	256	39,80	39,12	38,79	38,79	38,86	38,60	35,39
	512	43,84	43,19	42,95	42,98	42,99	42,78	39,75
	768	46,05	45,50	45,28	45,28	45,29	45,06	42,17
Soccer	256	34,17	33,19	32,76	32,94	32,92	32,80	31,17
	512	37,62	36,46	35,96	36,03	36,12	35,93	34,22
	768	39,82	38,65	38,10	38,17	38,24	37,98	36,12

Dla pozostałych sekwencji wizyjnych spadki wartości PSNR wynoszą dla sekwencji: Basket średnio 1,31dB, Harbour średnio 1,07dB, Ice średnio 1,31dB, Soccer średnio 1,73dB.

Dla wszystkich testowanych sekwencji wizyjnych wartości VSPD są tym większe, im częściej występuje zmiana sceny. Spadek wartości PSNR jest większy w przypadku sekwencji o płynnym ruchu i niskiej dynamice, np. City - w przeciwieństwie do sekwencji o wysokiej dynamice, np. Crew, Harbour. Wynika to z faktu, że w przypadku sekwencji wizyjnych o wysokiej dynamice, różnice między kolejnymi klatkami są mniejsze niż w sekwencjach o mniejszym ruchu. Odpowiednio, występowanie zmiany sceny w sekwencjach o wysokiej dynamice powoduje mniejszy spadek PSNR w porównaniu do sekwencji o mniejszej dynamice.

Godnym uwagi faktem jest znacznie większa wartość VSPD dla SC=1 w porównaniu do innych testowanych wartości SC. Wskaźnik VSPD jest od 2 do 4 razy wyższy (średnio 3,4 razy) dla SC=1 w porównaniu do SC = 2. Z drugiej strony wartości VSPD dla SC=4 w porównaniu do SC=2 wzrosły o 1,25 raza. Dla wszystkich testowanych sekwencji najniższą wartość VSPD uzyskano dla SC=32. Dla SC równego 16, 8 i 4 współczynnik VSPD zmienia się w stosunkowo małym zakresie. VSPD nieznacznie wzrasta przy SC=2 i bardzo silnie wzrasta przy SC=1. Oczywiście w praktyce zmiana treści sekwencji wizyjnej, co oznacza zmianę sceny dla każdego obrazu, może wystąpić bardzo rzadko. Niemniej w przypadku systemów monitoringu i nadzoru wizyjnego wykorzystywanych w IoT, największą wartość zwykle mają takie nagrania, gdzie występuje znacząca ilość ruchu, często w wyniku występowania jakiegoś zdarzenia objętego monitoringiem.

Przedstawione wyniki pokazują również, że w większości przypadków zmiana prędkości transmisji dla danej wartości SC nie powoduje znaczącej zmiany wartości VSPD. Ta sytuacja występuje szczególnie przy wyższych prędkościach transmisji. Wartość VSPD pozostaje na podobnym poziomie.

Tylko w nielicznych przypadkach wartość VSPD zależy wyraźnie od prędkości transmisji. Takie sytuacje dotyczą głównie przy SC=1 i sekwencji: Basket, City, Harbour, Ice i Soccer. Wynika z tego, że prędkość transmisji ma stosunkowo niewielki wpływ na wartość VSPD przy poszczególnych wartościach SC.

4. PODSUMOWANIE

W artykule przedstawiono wyniki badań eksperymentalnych nad kodekiem wideo HEVC w symulowanych warunkach zmiany sceny. Badanie pokazuje, że zmiana sceny w sekwencji wizyjnej prowadzi do spadku wartości PSNR. Im częstsza jest zmiana sceny, tym większy spadek PSNR. W systemach monitoringu i nadzoru wizyjnym stosowanych w urządzeniach Internetu rzeczy (IoT), często dopiero w momencie wystąpienia ruchu nagranie wizyjne niesie istotną wartość, kiedy dochodzi do jakiegoś, często nagłego zdarzenia. Znaczna ilość ruchu zaś powoduje negatywne skutki, jak przy zmianie sceny opisanej w artykule. Badania wskazują, że w takiej sytuacji jakość nagrania może się istotnie pogorszyć. Aby temu przeciwdziałać, warto rozważyć stosowanie rozwiązań umożliwiających zwiększenie prędkości transmisji w momencie wystąpienia znacznej ilości ruchu w nagraniach. Takie podejście umożliwi skompensowanie chwilowej utraty jakości spowodowanej ruchem, poprzez czasowe podniesienie prędkości transmisji rejestrowanych obrazów.

LITERATURA

1. R. Pereira R., Pereira E. G.: Video Streaming Considerations for Internet of Things, International Conference on Future Internet of Things and Cloud, Barcelona, 2014.
2. Kokkonis, G., Psannis, K.E., Roumeliotis, M. et al.: Real-time wireless multisensory smart surveillance with 3D-HEVC streams for internet-of-things (IoT). The Journal of Supercomputing 73, 2017.
3. ITU-T Rec. H.263, Video Coding for Low Bitrate Communication, 2005.
4. ISO/IEC 14496-10, International Standard, Generic Coding of Audio-Visual Objects, Part 10: advanced Video Coding, 6th ed., 2010, take ITU-T Rec. H.264, Edition 5.0 (version 11), 2010.
5. ISO/IEC International Standard 23008-2: ITU-T Recommendation H.265: April 2015, High efficiency video coding, 2015.
6. Sullivan, G.J., Ohm, J.-R., Han, W.-J., Wiegand, T.: Overview of the High Efficiency Video Coding (HEVC) Standard. IEEE Transactions on Circuits and Systems for Video Technology, 22 (12), 2012.
7. Wien M.: High Efficiency Video Coding: Coding Tools and Specification, Springer, 2015.
8. Ohm, J.R., Sullivan, G.J., Schwarz, H., Tan, T.K., Wiegand, T.: Comparison of the coding efficiency of video coding standards – Including High Efficiency Video Coding (HEVC). IEEE Transactions on Circuits and Systems for Video Technology, no. 22(12), 2102.
9. Huang, C. L., Liao, B. Y.: A robust scene-change detection method for video segmentation. IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 12, 2001.
10. Yi, X., Ling, N.: Fast pixel-based video scene change detection. IEEE International Symposium on Circuits and Systems, 2005.
11. Eom, Y., Park, S., Yoo, S., Choi, J. S., Cho, S.: An analysis of scene change detection in HEVC bitstream. In Semantic Computing (ICSC), IEEE International Conference on, 2015.
12. Test video sequences [Online] <ftp://ftp.tnt.uni-hannover.de/testsequences/>
13. HEVC test model reference software available online: <https://hevc.hhi.fraunhofer.de/>
14. Bossen, F.: Common test conditions and software reference configurations, Joint Collaborative Team on Video Coding (JCT-VC) of ITUT SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Doc. JCTVC-J1100, Stockholm, Sweden, 2012.